

QUANTITATIVE STABILITY OF VARIATIONAL SYSTEMS: I. THE EPIGRAPHICAL DISTANCE

HEDY ATTOUCH AND ROGER J.-B. WETS

ABSTRACT. This paper proposes a global measure for the distance between the elements of a variational system (parametrized families of optimization problems).

1. PRELIMINARIES

The study of the stability of the solutions of optimization problems is a central theme in the optimization literature. It has implications in model formulation, optimality characterizations, approximation theory (especially for infinite-dimensional problems), and in particular for numerical procedures. Most of the existing stability results are topological in nature, i.e., it is shown that under the appropriate conditions the minimum value function, or the set of optimal solutions, possesses some type of (semi)continuity. Although there are a few results of a quantitative nature, they are mostly limited to very specific transformations (perturbations) of a restricted class of problems. One of the reasons that there are essentially no “global” results is that there did not seem to exist a good metric, i.e., one with the appropriate theoretical properties and reasonably easy to compute, that could be used to measure the distance between two optimization problems.

In this paper, we study the *epi-distance* and show that it has many desirable properties. We then use it, in two subsequent papers [9, 10], to derive Hölderian and Lipschitzian properties for the optimal and ε -optimal solutions of optimization problems. The framework that serves as background to our study is that of *variational systems* as defined in Rockafellar and Wets [20], the stress being put on the *global* dependence of optimization problems on parameters that could affect the data that determines the objective as well as the constraints, even the structure of the problem itself.

Although optimization problems, in particular in infinite-dimensional spaces, have been our major motivation, one should point out that the results obtained for the epi-distance also have many implications in the convergence theory for operators. This theme is only developed to some extent in this paper, but the

Received by the editors June 6, 1989 and, in revised form, October 8, 1989.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 49B50, 26B25.

The second author's research was supported in part by the National Science Foundation.

© 1991 American Mathematical Society
0002-9947/91 \$1.00 + \$.25 per page

reader could get an idea of the possibilities from the observations made in §2 and the results in §5. Also, the results derived here in a functional framework have their counterparts for sets by specializing them to indicator functions. We illustrate this in just one case. In §3, we reformulate Theorem 3.7 in terms of sets. Similar types of corollaries could be worked out for most other theorems and propositions.

After the definition of the epi-distance in §1, §2 provides a useful criterion for the evaluation of the epi-distance in many practical situations. §3 makes a comparison between the epi-distance and other notions of distance based on epigraphical regularization (obtained with kernels of the type $(p\lambda)^{-1} \|\cdot\|^p$). §4 consists of a few basic observations about the topology induced by the epi-distance, and §5 considers the related convergence notion for monotone operators and relates the distance between the graphs of the subgradient mappings to the epi-distance. We conclude with a few results about an epigraphical calculus.

To begin with, let us review some notations and definitions. Unless specifically mentioned otherwise, we always denote by $(X, \|\cdot\|)$ a normed linear space and by d the distance function generated by the norm. For any subset C of X ,

$$d(x, C) := \inf_{y \in C} \|x - y\|$$

denotes the *distance from* x to C ; if $C = \emptyset$ we set $d(x, C) = \infty$. For any $\rho \geq 0$, $\rho\mathbb{B}$ denotes the ball of radius ρ and for any set C ,

$$C_\rho := C \cap \rho\mathbb{B}.$$

For $C, D \subset X$, the *excess* of C on D is defined as

$$e(C, D) := \sup_{x \in C} d(x, D),$$

with the (natural) convention that $e = 0$ if $C = \emptyset$. Note that the definition implies $e = \infty$ if C is nonempty and D is empty. For any $\rho \geq 0$, the ρ -(Hausdorff)-distance between C and D is given by

$$\text{haus}_\rho(C, D) = \max\{e(C_\rho, D), e(D_\rho, C)\}.$$

Definition 1.1. For $\rho \geq 0$, the ρ -epi-distance between two extended real valued functions f, g defined on X is

$$\text{haus}_\rho(f, g) := \text{haus}_\rho(\text{epi } f, \text{epi } g),$$

where the unit ball of $X \times \mathbb{R}$ is the set $\mathbb{B} := \mathbb{B}_{X \times \mathbb{R}} = \{(x, \alpha) : \|x\| \leq 1, |\alpha| \leq 1\}$.

One could trace this definition to the one used by Walkup and Wets [24] to measure the distance between convex cones, or that suggested by Mosco [18] to measure the distance between convex sets. But neither one of those earlier papers studies the properties of the epi-distance, or mentions its potential as a tool to obtain quantitative stability (convergence rates).

Proposition 1.2. *Let f_i ($i = 1, 2, 3$) be extended real valued functions defined on a normed linear space X . For any $\rho \geq 0$,*

- (i) *nonnegativity:* $\text{haus}_\rho(f_1, f_2) \geq 0$;
- (ii) *symmetry:* $\text{haus}_\rho(f_1, f_2) = \text{haus}_\rho(f_2, f_1)$;
- (iii) *triangle inequality:* for any $\rho > \inf_{\|x\| \leq \rho} f_i(x)$ ($i = 1, 2, 3$),

$$\text{haus}_\rho(f_1, f_3) \leq \text{haus}_{3\rho}(f_1, f_2) + \text{haus}_{3\rho}(f_2, f_3).$$

Moreover, if f_1 and f_2 are lower semicontinuous, then

- (iv) *for all $\rho > 0$, $\text{haus}_\rho(f_1, f_2) = 0$ if and only if $f_1 = f_2$.*

Note that the condition in (iii) is equivalent to $\rho > d(0, (\text{epi } f_i)_\rho)$.

Proof. Properties (i), (ii), and (iv) are self-evident. Proving (iii) is equivalent to showing that

$$\text{haus}_\rho(C_1, C_3) \leq \text{haus}_{3\rho}(C_1, C_2) + \text{haus}_{3\rho}(C_2, C_3),$$

where $C_i = \text{epi } f_i$ ($i = 1, 2, 3$) are subsets of the normed linear space $(X^\dagger = X \times \mathbb{R}, \|(x, \alpha)\|_\dagger := \max\{\|x\|, |\alpha|\})$. Let us prove that the above inequality holds with C_1, C_2, C_3 any subsets of a normed linear space Y . For $C, D \subset Y$ and $\rho \geq 0$, let

$$\delta_\rho(C, D) = \sup_{\|y\| \leq \rho} |d(y, C) - d(y, D)|,$$

where $\delta_\rho = \infty$ if either C and/or D is empty. Since $\rho\mathbb{B} \supset C_\rho$,

$$\delta_\rho(C, D) \geq \sup[d(y, D) : y \in C_\rho] = e(C_\rho, D),$$

and hence

$$(1.1) \quad \delta_\rho(C, D) \geq \text{haus}_\rho(C, D).$$

Conversely, for all $\rho > \max[\|y\|, d(0, C)]$,

$$d(y, C) \leq d(0, C) + \rho \leq 2\rho,$$

and thus $d(y, C) = d(y, C_{3\rho})$. It follows that

$$\sup\{d(y, D) - d(y, C) : \|y\| \leq \rho\} \leq \sup\{d(y, D) - d(y, C_{3\rho})\} \leq e(C_{3\rho}, D).$$

With the symmetric inequality, obtained when interchanging the roles of C and D , this becomes

$$(1.2) \quad \delta_\rho(C, D) \leq \text{haus}_{3\rho}(C, D).$$

Since δ_ρ clearly satisfies the triangle inequality, (1.1) and (1.2) imply

$$\begin{aligned} \text{haus}_\rho(C_1, C_3) &\leq \delta_\rho(C_1, C_3) \leq \delta_\rho(C_1, C_2) + \delta_\rho(C_2, C_3) \\ &\leq \text{haus}_{3\rho}(C_1, C_2) + \text{haus}_{3\rho}(C_2, C_3), \end{aligned}$$

provided $\rho > d(0, C_i)$ ($i = 1, 2, 3$). \square

Rather than defining the epi-distance as done here, one could have considered the Hausdorff distance between the intersection of both $\text{epi } g$ and $\text{epi } f$ with the ρ -ball, as done in Salinetti and Wets [21]. In general, this distance does not fill our needs, because it does not induce epi-convergence. However, in the convex case it would not matter, since it induces the same uniform structure as the epi-distance as we show next. We begin with a couple of preliminary results.

Lemma 1.3. *Let X be a normed linear space and C a closed convex set such that $C_{\rho_0} \neq \emptyset$. Then for any $\rho > \rho_0$ and $\eta \geq 0$*

$$\text{haus}(C_{\rho+\eta}, C_\rho) \leq [(\rho + \rho_0)/(\rho - \rho_0)] \cdot \eta,$$

which implies that the map $\eta \mapsto \text{haus}(C_{\rho+\eta}, C_\rho)$ is Lipschitzian on \mathbb{R}_+ .

Proof. The argument is based on duality. From Hörmander's classical formula (see [6, §3], for example),

$$\text{haus}(C_{\rho+\eta}, C_\rho) = \sup\{|s(C_{\rho+\eta}, x^*) - s(C_\rho, x^*)| : \|x^*\| \leq 1\},$$

when $s(D, x^*) = \sup\{\langle x^*, x \rangle : x \in D\}$ is the support function of D . Note that $s(C_\rho, \cdot) = (\delta_C + \delta_{\rho\mathbb{B}})^*$, where δ_C is the indicator function of the set C . Moreover, $\delta_{\rho\mathbb{B}}$ is continuous at a point of the domain of δ_C —because C_{ρ_0} is nonempty and $\rho > \rho_0$ —which means that

$$(1.3) \quad s(C_\rho, x^*) = \min\{s(C, y^*) + \rho\|x^* - y^*\| : y^* \in X^*\}$$

with the minimum attained at some point y_ρ^* . Thus,

$$\begin{aligned} s(C_{\rho+\eta}, x^*) - s(C_\rho, x^*) & \\ & \leq \{s(C, y_\rho^*) + (\rho + \eta)\|x^* - y_\rho^*\|\} - \{s(C, y_\rho^*) + \rho\|x^* - y_\rho^*\|\} \\ & = \eta \cdot \|x^* - y_\rho^*\|. \end{aligned}$$

The proof is completed by showing that $\|x^* - y_\rho^*\| \leq (\rho + \rho_0)(\rho - \rho_0)^{-1}\|x^*\|$. Indeed,

$$(1.4) \quad \rho\|y_\rho^*\| - \rho\|x^*\| \leq \rho\|x^* - y_\rho^*\| \leq \rho\|x^*\| + \rho_0\|y_\rho^*\|,$$

where the last inequality from (1.3) with the observations

$$\begin{aligned} s(C_\rho, x^*) & \leq s(\rho\mathbb{B}, x^*) = \rho\|x^*\|, \\ s(C, y_\rho^*) & \leq \langle y_\rho^*, x_0 \rangle \geq -\rho_0\|y_\rho^*\|, \end{aligned}$$

with x_0 any point in C_{ρ_0} . Thus, $\|y_\rho^*\| \leq 2\rho(\rho - \rho_0)^{-1}\|x^*\|$, which combined with the last relation in (1.4) yields the asserted inequality; recall that $\|x^*\| \leq 1$. \square

In connection with this lemma, let us observe that $(\rho + \rho_0)/(\rho - \rho_0)$ the constant used to bound the Hausdorff distance between $C_{\rho+\eta}$ and C_ρ goes to

1 (resp. ∞) as ρ goes to ∞ (resp. ρ_0) is not necessarily the best possible constant that could be obtained to obtain this estimate. The best constant depends on the geometry of the space $(X, \|\cdot\|)$. For example, if X is a Hilbert space, we can replace $(\rho + \rho_0)/(\rho - \rho_0)$ by $[(\rho + \rho_0)/(\rho - \rho_0)]^{1/2}$. To see this, we calculate a bound for $e(C_{\rho+\eta}, C_\rho)$. Let p_0 be the projection of the origin on the set C . Then $\|p_0\| \leq \rho_0$ since $C_{\rho_0} \neq \emptyset$. Pick $x \in C_{\rho+\eta} \setminus C_\rho$ and let $L = \{p_0 + \lambda(x - p_0) | \lambda \geq 0\}$ be the half-line rooted at p_0 and passing through x . We are looking for an upper bound on $d(x, C_\rho)$ and to do this we can restrict ourselves to the plane that includes L and $\{0\}$. With $y \in L \cap \text{bdy} \rho \mathbb{B}$ and $z \in L \cap (\text{bdy}(\rho + \eta) \mathbb{B})$, $d(x, C_\rho) \leq \|y - z\|$. Because of the Hilbert geometry of the space, $\|y - z\|$ is maximal when the angle between the line that includes L and the line passing through 0 and p_0 is $\pi/2$. From the Pythagorean Theorem it follows that $e(C_{\rho+\eta}, C_\rho) \leq \sqrt{(\rho + \eta)^2 - \rho_0^2} - \sqrt{\rho^2 - \rho_0^2}$. A simple calculation now shows that this quantity is bounded above by $[(\rho + \rho_0)/(\rho - \rho_0)]^{1/2} \eta$, given that $\eta \leq 1$.

Proposition 1.4. *Let C and D be two closed subsets of a normed linear space X such that C_{ρ_0} and D_{ρ_0} are nonempty. Then for all $\rho \geq \rho_0$*

$$\text{haus}_\rho(C, D) \leq \text{haus}(C_\rho, D_\rho).$$

Moreover, if C and D are convex, then for $\rho > \rho_0$

$$\text{haus}(C_\rho, D_\rho) \leq \frac{2\rho}{\rho - \rho_0} \text{haus}_\rho(C, D).$$

Proof. The first inequality is self-evident. The second one follows from the “triangle inequality” for the excess function, $e(R, T) \leq e(R, S) + e(S, T)$, for any sets R , S , and T . It implies that $\text{haus}(C_\rho, D_\rho) \leq \beta_1 + \beta_2$, where for any $\eta > 0$,

$$\begin{aligned} \beta_1 &:= \max[e(D_\rho, C_{\rho+\eta}), e(C_\rho, D_{\rho+\eta})], \\ \beta_2 &:= \max[e(C_{\rho+\eta}, C_\rho), e(D_{\rho+\eta}, D_\rho)]. \end{aligned}$$

When C and D are convex, we use Lemma 1.3 to obtain

$$\beta_2 \leq [(\rho + \rho_0)/(\rho - \rho_0)] \eta.$$

With $\eta = \text{haus}_\rho(C, D)$,

$$e(C_\rho, D_{\rho+\eta}) = e(C_\rho, D) \quad \text{and} \quad e(D_\rho, C_{\rho+\eta}) = e(D_\rho, C),$$

i.e., $\beta_1 = \text{haus}_\rho(C, D)$. This, with the preceding bound for β_2 , yields the estimate. \square

Corollary 1.5. *Let X be a normed linear space and f and g two proper extended real valued lower semicontinuous functions defined on X . Let $\rho_0 > 0$ be such that $(\text{epi } f)_{\rho_0}$ and $(\text{epi } g)_{\rho_0}$ are nonempty. Then for all $\rho \geq \rho_0$,*

$$\text{haus}_\rho(f, g) \leq \text{haus}((\text{epi } f)_\rho, (\text{epi } g)_\rho).$$

Moreover, if the functions are also convex, for $\rho > \rho_0$

$$\text{haus}((\text{epi } f)_\rho, (\text{epi } g)_\rho) \leq \frac{2\rho}{\rho - \rho_0} \text{haus}_\rho(f, g).$$

Proof. Simply apply the proposition to the closed epigraphs of f and g . \square

2. THE KENMOCHI CONDITIONS

The Kenmochi conditions provide a practical criterion for computing, or at least estimating, the epi-distance between two functions.

Theorem 2.1. *Suppose f, g are proper extended real valued functions defined on a normed linear space X , both minorized by $-\alpha_0\|\cdot\|^p - \alpha_1$ for some $\alpha_0 \geq 0$, $\alpha_1 \in \mathbb{R}$, and $p \geq 1$. Let $\rho_0 > 0$ be such that $(\text{epi } f)_{\rho_0}$ and $(\text{epi } g)_{\rho_0}$ are nonempty.*

(a) *Then the following conditions—to be called the Kenmochi conditions—hold: for all $\rho > \rho_0$ and $x \in \text{dom } f$ such that $\|x\| \leq \rho$, $|f(x)| \leq \rho$, for every $\varepsilon > 0$ there exists some $\tilde{x}_\varepsilon \in \text{dom } g$ that satisfies*

$$(2.1) \quad \|x - \tilde{x}_\varepsilon\| \leq \text{haus}_\rho(f, g) + \varepsilon, \quad g(\tilde{x}_\varepsilon) \leq f(x) + \text{haus}_\rho(f, g) + \varepsilon$$

as well as a symmetric condition with the roles of f and g interchanged.

(b) *Conversely, assuming that for all $\rho > \rho_0 > 0$ there exists a “constant” $\eta(\rho) \in \mathbb{R}_+$, depending on ρ , such that for all $x \in \text{dom } f$ with $\|x\| \leq \rho$, $|f(x)| \leq \rho$, there exists $\tilde{x} \in \text{dom } g$ that satisfies*

$$(2.2) \quad \|x - \tilde{x}\| \leq \eta(\rho), \quad g(\tilde{x}) \leq f(x) + \eta(\rho),$$

and the symmetric condition (interchanging the roles of f and g), then with $\rho_1 := \rho + \alpha_0\rho^p + |\alpha_1|$,

$$(2.3) \quad \text{haus}_\rho(f, g) \leq \eta(\rho_1).$$

Proof. It suffices to observe that

(i) $\text{haus}_\rho(\text{epi } f, \text{epi } g) \leq \theta$ if and only if for every $\varepsilon > 0$

$$(\text{epi } f)_\rho \subset \text{epi } g + (\theta + \varepsilon)\mathbb{B} \quad \text{and} \quad (\text{epi } g)_\rho \subset \text{epi } f + (\theta + \varepsilon)\mathbb{B};$$

(ii) these inclusions yield exactly the Kenmochi conditions (2.1) if one remembers that $\text{epi } g$ is an epigraph;

(iii) estimate (2.3) is obtained by calculating an upper bound for θ in terms of ρ and $\eta(\rho)$. That is done next.

Given any $(x, \mu) \in (\text{epi } f)_\rho$, i.e., $\|x\| \leq \rho$, $|\mu| \leq \rho$, $\mu \geq f(x)$, we have that $|f(x)| \leq \rho_1$. By (2.2) there exists some $\tilde{x} \in \text{dom } g$ with $\|x - \tilde{x}\| \leq \eta(\rho_1)$ such that

$$g(\tilde{x}) \leq f(x) + \eta(\rho_1) \leq \mu + \eta(\rho_1).$$

There remains only to observe that if $g(\tilde{x}) \geq \mu$, then $|\mu - g(\tilde{x})| = g(\tilde{x}) - \mu \leq \eta(\rho_1)$, and

$$d((x, \mu), \text{epi } g) \leq d((x, \mu), (\tilde{x}, g(\tilde{x}))) \leq \eta(\rho_1).$$

On the other hand, if $\mu \geq g(\tilde{x})$, then $(\tilde{x}, \mu) \in \text{epi } g$ and consequently

$$d((x, \mu), \text{epi } g) \leq d((x, \mu), (\tilde{x}, \mu)) \leq \eta(\rho_1). \quad \square$$

Remark 2.2. Theorem 2.1 tells us that in order to compute $\text{haus}_\rho(f, g)$ we have to find the best constant $\eta(\rho)$ for which the condition (2.2) holds. This condition had been introduced by Kenmochi [16] (see also Attouch and Damlamian [4]) to study the existence of strong solutions for evolution problems of the following type:

$$0 \in \frac{du}{dt} + \partial f(t, u(t)); \quad u(0) = u_0.$$

The time dependence of f with respect to t , in our terminology, can now be expressed as an absolute continuity property of the map $t \mapsto f(t)$. It can be formulated as follows: there exist $b \in C([0, T]; H) \cap W^{1,2}([0, T]; H)$ and a , an increasing function, such that:

$$\forall 0 \leq s \leq t \leq T, \forall x \in \text{dom } f(s, \cdot), \exists \tilde{x} \in \text{dom } f(t, \cdot) \text{ such that}$$

$$\|x - \tilde{x}\| \leq |b(t) - b(s)| \cdot (1 + \|x\|),$$

$$f(t, \tilde{x}) \leq f(s, x) + (a(t) - a(s))(|f(s, x)| + \|x\|^2 + 1).$$

Thus, $\forall x \in \text{dom } f(s, \cdot)$ with $\|x\| \leq \rho$, $|f(s, \cdot)| \leq \rho$, we have the existence of some $\tilde{x} \in \text{dom } f(t, \cdot)$ such that

$$\|x - \tilde{x}\| \leq (1 + \rho)|b(t) - b(s)|,$$

$$f(t, \tilde{x}) \leq f(s, x) + (1 + \rho + \rho^2)(a(t) - a(s)).$$

Taking $\eta(\rho) = \max\{(1 + \rho)|b(t) - b(s)|, (1 + \rho + \rho^2)(a(t) - a(s))\}$, we see that condition (2.2) is satisfied. \square

3. COMPARISON WITH THE $d_{\lambda, \rho}$ -DISTANCES

This section is devoted to the relationship between the epi-distance and the distances introduced in Attouch and Wets [6, 7], based on epigraphical regularizations. Although one can envisage more general kernels (see Wets [25] and Attouch, Azé, and Wets [3, Propositions 3.1 and 3.2]), for simplicity's sake we shall restrict ourselves to regularizations with respect to *kernels* $k: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ of the type

$$k(r) = \frac{1}{p} r^p \quad \text{for some } 1 \leq p < \infty.$$

The *epigraphical regularization* f_λ of parameter $\lambda > 0$ of a function $f: X \rightarrow \overline{\mathbb{R}}$ (with X a normed linear space) is defined by

$$f_\lambda := f +^e \lambda^{-1} k(\|\cdot\|),$$

where $+^e$ denotes *epigraphical sum* (inf-convolution):

$$f_\lambda(x) = \inf_{u \in X} \left\{ f(u) + \frac{1}{\lambda p} \|x - u\|^p \right\}.$$

With $p = 1$, f_λ is called the Baire-Wijsman approximate, and with $p = 2$, the Moreau-Yosida approximate of f (cf. [6]). Assume that $f + \alpha(\|\cdot\|^p + 1) \geq 0$ for some $\alpha > 0$, then for $0 < \lambda < (\alpha p)^{-1} 2^{1-p}$, f_λ is a continuous locally Lipschitz function on X , as will be shown below.

Now fixing the parameter p in the kernel k once and for all, we can define the following distance between two functions f and g :

$$d_{\lambda,p}(f, g) = \sup_{\|x\| \leq \rho} |f_\lambda(x) - g_\lambda(x)|.$$

Assuming that f and g are proper, this quantity is well defined since both f_λ and g_λ are then bounded on bounded sets. These distance functions induce epi-convergence, and in [6, Theorem 2.33 and Corollary 2.42], we obtain quantitative stability results in terms of the resolvents of the Moreau-Yosida approximations.

We start with some basic properties of f_λ .

Lemma 3.1. *Let $p \in [1, \infty)$, $\alpha_0 \geq 0$, and $\alpha_1 \in \mathbb{R}$. Suppose $f \not\equiv \infty$ is an extended real valued function defined on $(X, \|\cdot\|)$ such that $f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0$. Then f_λ is finite valued for all $\lambda \in (0, (\alpha_0 p)^{-1} 2^{1-p})$. Moreover, for all $x_0 \in X$ and $\beta \in \mathbb{R}$*

$$f_\lambda(x + x_0) + \beta = (f(\cdot + x_0) + \beta)_\lambda(x).$$

Proof. The inequality

$$f_\lambda(x) \geq \inf_u [-\alpha_0 \|x - u\|^p + (p\lambda)^{-1} \|u\|^p] - \alpha_1 \geq -\alpha_0 2^{p-1} \|x\|^p - \alpha_1$$

follows from $\|x - u\|^p \leq 2^{p-1}(\|u\|^p + \|x\|^p)$ and $\lambda < (\alpha_0 p)^{-1} 2^{1-p}$. For an upper bound, let x_0 be such that $f(x_0)$ is finite; then

$$f_\lambda(x) \leq f(x_0) + (p\lambda)^{-1} \|x - x_0\|^p.$$

Finally,

$$\begin{aligned} (f(\cdot + x_0) + \beta)_\lambda(x) &= \inf_u [f(u + x_0) + \beta + (p\lambda)^{-1} \|x - u\|^p] \\ &= \inf [f(v) + (p\lambda)^{-1} \|x + x_0 - v\|^p] + \beta. \quad \square \end{aligned}$$

The next lemma extends Theorems 2.64 of Attouch [1] to epigraphical regularizations involving any kernel of the type $(\lambda p)^{-1} \|\cdot\|^p$ for $p \geq 1$.

Lemma 3.2. *Let $f \not\equiv \infty$ be an extended real valued function defined on $(X, \|\cdot\|)$ such that for some $\alpha_0 \geq 0$ and $\alpha_1 \in \mathbb{R}$, $f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0$ for given $p \geq 1$. Then for any $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$,*

$$f_\lambda = f + {}^c (p\lambda)^{-1} \|\cdot\|^p$$

is locally Lipschitz, i.e.,

$$|f_\lambda(x) - f_\lambda(y)| \leq \lambda^{-1} \kappa \|x - y\|,$$

where the Lipschitz constant κ depends continuously on $\|x\|$, $\|x - y\|$, α_0 , λ , and p ; it depends on f only through the value $f(x_f)$ and the norm $\|x_f\|$ at a point x_f at which f is finite.

Proof. We have already established that under these assumptions f_λ is finite valued. To simplify the calculations, let us first suppose that $f(0) = 0$. Now from the definition of f_λ , it follows that for all $x \in X$ and $\varepsilon > 0$, there exists u_x^ε such that

$$f_\lambda(x) \leq f(u_x^\varepsilon) + (p\lambda)^{-1} \|x - u_x^\varepsilon\|^p \leq f_\lambda(x) + \varepsilon,$$

and thus, since f is minorized by $-\alpha_0 \|\cdot\|^p - \alpha_1$,

$$-\alpha_0 \|u_x^\varepsilon\|^p - \alpha_1 + (p\lambda)^{-1} \|x - u_x^\varepsilon\|^p \leq f_\lambda(x) + \varepsilon \leq (p\lambda)^{-1} \|x\|^p + \varepsilon,$$

where the last inequality comes from the upper bound

$$f_\lambda(x) \leq f(0) + (p\lambda)^{-1} \|x - 0\|^p = (p\lambda)^{-1} \|x\|^p.$$

Since $\|u_x^\varepsilon\|^p \leq 2^{p-1} (\|x - u_x^\varepsilon\|^p + \|x\|^p)$, it follows that

$$\|x - u_x^\varepsilon\|^p ((p\lambda)^{-1} - \alpha_0 2^{p-1}) \leq ((p\lambda)^{-1} + \alpha_0 2^{p-1}) \|x\|^p + \alpha_1 + \varepsilon.$$

With $\alpha' := \alpha_0 2^{p-1}$, this yields

$$\|x - u_x^\varepsilon\|^p \leq \frac{1 + \alpha' \lambda}{1 - \alpha' \lambda} \|x\|^p + \frac{p\lambda}{1 - \alpha' \lambda} (\alpha_1 + \varepsilon).$$

For any $y \in X$, we have

$$\begin{aligned} f_\lambda(y) &\leq f(u_x^\varepsilon) + (p\lambda)^{-1} \|y - u_x^\varepsilon\|^p \\ &\leq f_\lambda(x) + \varepsilon + \lambda^{-1} \left[\frac{1}{p} \|y - u_x^\varepsilon\|^p - \frac{1}{p} \|x - u_x^\varepsilon\|^p \right]. \end{aligned}$$

we use the convexity of $t \mapsto p^{-1} t^p$ on \mathbb{R}_+ , and the subgradient inequality to obtain

$$\frac{1}{p} (\|y - x\| + \|x - u_x^\varepsilon\|)^p - \frac{1}{p} \|x - u_x^\varepsilon\|^p \leq (\|y - x\| + \|x - u_x^\varepsilon\|)^{p-1} \|x - y\|,$$

and since $\|y - x\| + \|x - u_x^\varepsilon\| \geq \|y - u_x^\varepsilon\|$, it follows that

$$f_\lambda(y) - f_\lambda(x) \leq \varepsilon + \lambda^{-1} \|y - x\| (\|y - x\| + \|x - u_x^\varepsilon\|)^{p-1}.$$

We now use the estimate we have for $\|x - u_x^\varepsilon\|$ and let ε go to zero. This yields

$$f_\lambda(y) - f_\lambda(x) \leq \lambda^{-1} \kappa_x \|y - x\|,$$

where

$$\kappa_x = \left(\|y - x\| + \left[\frac{1 + \alpha' \lambda}{1 - \alpha' \lambda} \|x\|^p + \frac{p\lambda}{1 - \alpha' \lambda} \alpha_1 \right]^{1/p} \right)^{p-1}$$

is a “constant” that depends only on $\|x\|$, $\|y - x\|$, α_0 , p , λ . Interchanging the roles of x and y in the above, we obtain a similar inequality with a constant κ_y . Setting $\kappa = \max[\kappa_x, \kappa_y]$ yields the desired inequality.

If $f(0) \neq 0$, let x_0 be such that $f(x_0) \in \mathbb{R}$. Then $f(\cdot + x_0) - f(x_0)$ is a function that takes on the value 0 at $x = 0$, and moreover (cf. Lemma 3.1),

$$(f(\cdot + x_0) - f(x_0))_\lambda = f_\lambda(\cdot + x_0) - f(x_0).$$

From our earlier argument and this last identity, it follows that

$$|f_\lambda(y) - f_\lambda(x)| \leq \frac{1}{\lambda} \kappa \|y - x\|,$$

where in the definition of κ_x the term $\|x\|^p$ is replaced by $\|x - x_0\|^p$ and similarly in κ_y and α_1 is replaced by $\alpha_1 - f(x_0)$. \square

Lemma 3.3. *Let X be a normed linear space, and f and g two extended real valued proper functions defined on X such that for some $\alpha_0 \geq 0$ and $\alpha_1 \in \mathbb{R}$,*

$$f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0, \quad g + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0.$$

Then for $1 \leq p < \infty$, any $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$, and $\rho \geq \max[f_\lambda(0), g_\lambda(0)]$, we have

$$(3.1) \quad \text{haus}_\rho(f_\lambda, g_\lambda) \leq \text{haus}_\gamma(f, g),$$

where the constant γ , that depends on ρ , is defined by (3.3).

Proof. There is nothing to prove if $\text{haus}_\gamma(f, g) = \infty$, so let us assume that $\text{haus}_\gamma(f, g)$ is finite; note also that f, g proper implies that $\text{epi } f$ and $\text{epi } g$ are nonempty, and that $\rho > \max[f_\lambda(0), g_\lambda(0)]$ implies that $\text{haus}_\rho(f_\lambda, g_\lambda)$ is finite. To have $\text{haus}_\gamma(f, g) < \eta$ means that

$$(\text{epi } f)_\gamma \subset \eta(\text{epi } g) \quad \text{and} \quad (\text{epi } g)_\gamma \subset \eta(\text{epi } f),$$

where $\eta D := \{x | d(x, D) \leq \eta\}$ is the η -fattening of D . From this, it follows

$$(\text{epi } f)_\gamma + \text{epi}(p\lambda)^{-1} \|\cdot\|^p \subset \eta(\text{epi } g) + \text{epi}(p\lambda)^{-1} \|\cdot\|^p,$$

and this inclusion, with $\text{epi } g + \text{epi}(\lambda p)^{-1} \|\cdot\|^p \subset \text{epi } g_\lambda$, yields $(\text{epi } f)_\gamma + \text{epi}(p\lambda)^{-1} \|\cdot\|^p \subset \eta(\text{epi } g_\lambda)$. Since

$$\text{epi}_s f + \text{epi}_s(p\lambda)^{-1} \|\cdot\|^p = \text{epi}_s f_\lambda,$$

where $\text{epi}_s h := \{(x, \alpha) | \alpha > f(x)\}$ is the *strict epigraph* of h , it suffices to prove that γ can be chosen so that

$$(3.2) \quad (\text{epi}_s f + \text{epi}_s(\lambda p)^{-1} \|\cdot\|^p)_\rho \subset (\text{epi } f)_\gamma + \text{epi}(\lambda p)^{-1} \|\cdot\|^p.$$

Indeed, the last three identities would imply $e((\text{epi}_s f_\lambda)_\rho, \text{epi } g_\lambda) \leq \eta$, or still, for all $\varepsilon > 0$ and for all $\eta > \text{haus}_\gamma(f, g)$, $e((\text{epi } f_\lambda)_{\rho-\varepsilon}, \text{epi } g_\lambda) \leq \eta + \varepsilon$. The asserted inequality (3.1) now follows from the fact that f_λ is locally Lipschitz (Lemma 3.2), and that f and g have symmetric roles.

We turn to (3.2). Let $(x, \alpha) \in (\text{epi}_s f_\lambda)_\rho$. By definition of f_λ , there exists $u_x \in X$ such that $\alpha > f(u_x) + (p\lambda)^{-1} \|x - u_x\|^p$. Moreover, since $(x, \alpha) = (u_x, f(u_x)) + (x - u_x, \alpha - f(u_x))$ and $\alpha - f(u_x) \geq (p\lambda)^{-1} \|x - u_x\|^p$, it suffices

to show that there exists γ such that $\|u_x\| \leq \gamma$ and $|f(u_x)| \leq \gamma$. From $|\alpha| \leq \rho$ and $f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0$, it follows that

$$-\alpha_0 \|u_x\|^p - \alpha_1 + (p\lambda)^{-1} \|x - u_x\|^p \leq \rho.$$

The same calculation as in Lemma 3.2 yields

$$\|u_x\|^p \leq (2^{1-p} - \alpha_0 \lambda p)^{-1} (\rho^p + p\lambda\rho + p\lambda\alpha_1) := \gamma_1(\rho).$$

From $\alpha > f(u_x)$, we obtain $|f(u_x)| \leq \sup\{\rho; \alpha_0 \|u_x\|^p + \alpha_1\}$, and thus we can define γ as

$$(3.3) \quad \gamma = \gamma(\rho) := \sup\{\rho; \gamma_1 \rho^{1/p}; \alpha_0 \gamma_1(\rho) + \alpha_1\}. \quad \square$$

Theorem 3.4. Suppose f and g are two extended real valued functions defined on a normed linear space $(X, \|\cdot\|)$ such that $f \geq -\alpha_0 \|\cdot\|^p - \alpha_1$ and $g \geq -\alpha_0 \|\cdot\|^p - \alpha_1$ for some $p \in [1, \infty)$, $\alpha_0 \geq 0$, and $\alpha_1 \in \mathbb{R}$. Then for all $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$ and $\rho \geq 0$,

$$d_{\lambda, \rho}(f, g) \leq \beta(\lambda, \rho) \text{haus}_{\gamma(\lambda, \rho)}(f, g),$$

with the “constants” γ and β , in the case when f and g are not identically ∞ , depending only on ρ , α_0 , α_1 , λ , p , and $f(x_f)$, $g(x_g)$, $\|x_f\|$, $\|x_g\|$, where x_f and x_g are arbitrary points in $\text{dom } f$ and $\text{dom } g$.

Proof. Excluding the cases $f \equiv \infty$ or/and $g \equiv \infty$, when the inequality is trivially satisfied, the functions f_λ and g_λ are finite valued, in fact equi-locally Lipschitz (cf. Lemma 3.2). This can be used to conclude that whenever $\|x\| \leq \rho$, both $f_\lambda(x)$ and $g_\lambda(x)$ are bounded in absolute value by

$$\rho' \geq \max\{|f_\lambda(0) \pm \tfrac{1}{\lambda} \kappa_\rho \rho|, |g_\lambda(0) \pm \tfrac{1}{\lambda} \kappa_\rho \rho|\},$$

where κ_ρ is the Lipschitz constant calculated in Lemma 3.2 that depends on ρ , α_0 , λ , p , $f(x_f)$, $g(x_g)$, $\|x_f\|$, $\|x_g\|$. Note that also $f_\lambda(0)$ can be bounded by constants that depend only on α_0 , α_1 , the norm of x_f , and analogously for $g_\lambda(0)$. Setting $\rho_1(\lambda, \rho) = \rho_1 := \max[\rho, \rho']$, let us estimate $g_\lambda(x) - f_\lambda(x)$ when $\|x\| \leq \rho$. By the above and Kenmochi's conditions (2.1), for all $\varepsilon > 0$ there exists y such that

$$\|x - y\| \leq \text{haus}_{\rho_1}(f_\lambda, g_\lambda) + \varepsilon, \quad g_\lambda(y) \leq f_\lambda(x) + \text{haus}_{\rho_1}(f_\lambda, g_\lambda) + \varepsilon,$$

since $\|x\| \leq \rho_1$ and $|f_\lambda(x)| \leq \rho_1$. Because $g_\lambda(x) - f_\lambda(x) = (g_\lambda(x) - g_\lambda(y)) + (g_\lambda(y) - f_\lambda(x))$, it follows that

$$\begin{aligned} g_\lambda(x) - f_\lambda(x) &\leq (p\lambda)^{-1} \kappa_{\rho_2} \|x - y\| + \text{haus}_{\rho_1}(f_\lambda, g_\lambda) + \varepsilon \\ &\leq \beta \text{haus}_{\rho_1}(f_\lambda, g_\lambda) + \beta \varepsilon, \end{aligned}$$

where $\beta := \beta(\lambda, \rho) = (p\lambda)^{-1} \kappa_{\rho_2} + 1$ and $\rho_2 := \rho + \text{haus}_{\rho_1}(f_\lambda, g_\lambda) + 1$; κ_{ρ_2} is the Lipschitz constant coming from Lemma 3.2, valid for both f and g , when $x = 0$ and $\|y\| \leq \rho_2$. All these quantities depend only on ρ , α_0 , α_1 , λ , p ,

$f(x_f)$, $g(x_g)$, $\|x_f\|$, and $\|x_g\|$. With an analogous inequality obtained when the roles of f and g are interchanged, and after letting ε go to 0, this yields

$$d_{\lambda, \rho}(f, g) = \sup_{\|x\| \leq \rho} |f_{\lambda}(x) - g_{\lambda}(x)| \leq \beta \text{haus}_{\rho_1}(f_{\lambda}, g_{\lambda}) \leq \beta \text{haus}_{\gamma}(f, g),$$

where the last inequality follows from Lemma 3.3; the constant γ is that obtained when replacing ρ by ρ_1 in formula (3.3). \square

The arguments that we have used in the proof of Theorem 3.4 are geometric in nature, we give another proof in the Appendix that is of analytic type. It yields a more direct calculation of the Lipschitz constant, but does not explicitly bring to the fore the properties of regularized functions (Lemmas 3.1 and 3.2), and the useful inequality in Lemma 3.3.

Our next task is to derive an appropriate bound for haus_{ρ} in terms of $d_{\lambda, \rho}$. Again we start with some preparatory lemmas that are of independent interest.

Lemma 3.5. *Let X be a normed space. Suppose f and g are proper, extended real valued functions defined on X . Then for all $\lambda > 0$ and $\rho \geq 0$,*

$$\text{haus}_{\rho}(f_{\lambda}, g_{\lambda}) \leq d_{\lambda, \rho}(f, g).$$

Proof. If $\mu \geq g_{\lambda}(x)$, $\|x\| \leq \rho$, and $|\mu| \leq \rho$, then $\mu \geq f_{\lambda}(x) - d_{\lambda, \rho}(f, g)$. This implies that $(\mu + d_{\lambda, \rho}(f, g), x) \in \text{epi } f_{\lambda}$, and thus

$$e((\text{epi } g_{\lambda})_{\rho}, \text{epi } f_{\lambda}) \leq d_{\lambda, \rho}(f, g). \quad \square$$

Lemma 3.6. *Let X be a normed linear space, and f a proper extended real valued function defined on X , minorized by $-\alpha_0 \|\cdot\|^p - \alpha_1$ for some $p \geq 1$, $\alpha_0 \geq 0$, and $\alpha_1 \in \mathbb{R}$. Then, for any $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$ and $\rho \geq 0$,*

$$\text{haus}_{\rho}(f_{\lambda}, f) \leq \lambda^{1/p} \left[\frac{p(\alpha_0 2^{p-1} \rho^p + \rho + \alpha_1)}{1 - \alpha_0 p \lambda 2^{p-1}} \right]^{1/p}.$$

Proof. Since $f_{\lambda} \leq f$, $e((\text{epi } f)_{\rho}, \text{epi } f_{\lambda}) = 0$. To calculate an upper bound for $e((\text{epi } f_{\lambda})_{\rho}, \text{epi } f)$, let $(x, \mu) \in (\text{epi } f_{\lambda})_{\rho}$, with $\|x\| \leq \rho$, $|\mu| \leq \rho$, and $\mu \geq f_{\lambda}(x)$, and denote by u_x^{ε} an element such that $\varepsilon + \mu \geq f(u_x^{\varepsilon}) + (p\lambda)^{-1} \|x - u_x^{\varepsilon}\|^p$. Note that $(u_x^{\varepsilon}, \varepsilon + \mu) \in \text{epi } f$, and thus

$$d((x, \mu), \text{epi } f) \leq \max[\|x - u_x^{\varepsilon}\|, |\varepsilon + \mu - \mu|].$$

It thus suffices to obtain a bound on $\|x - u_x^{\varepsilon}\|$. From the minorization of f , and $|\mu| \leq \rho$, it follows that

$$\varepsilon + \rho \geq -\alpha_0 \|u_x^{\varepsilon}\|^p - \alpha_1 + \frac{1}{p\lambda} \|x - u_x^{\varepsilon}\|^p.$$

We rely on the inequalities $\|u_x^{\varepsilon}\|^p \leq 2^{p-1}(\|x\|^p + \|u_x^{\varepsilon} - x\|^p)$, $\|x\| \leq \rho$, and $(\lambda p)^{-1} - \alpha_0 2^{p-1} > 0$, to obtain

$$\|x - u_x^{\varepsilon}\|^p \leq ((\lambda p)^{-1} - \alpha_0 2^{p-1})^{-1} (\alpha_0 2^{p-1} \rho^p + \rho + \alpha_1),$$

which yields the asserted bound. \square

Theorem 3.7. *Suppose X is a normed linear space and f and g are proper, extended real valued functions defined on X such that for given $p \geq 1$, and some $\alpha_0 \geq 0$, $\alpha_1 \in \mathbb{R}$,*

$$f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0 \quad \text{and} \quad g + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0.$$

Then, for all λ satisfying $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$ and

$$\rho \geq \max[d(0, \text{epi } f), d(0, \text{epi } g)],$$

we have

$$\text{haus}_\rho(f, g) \leq d_{\lambda, \rho}(f, g) + 2(\lambda p)^{1/p} \left[\frac{\alpha_0 2^{p-1} (9\rho)^p + 9\rho + \alpha_1}{1 - \alpha_0 p \lambda 2^{p-1}} \right]^{1/p}.$$

Proof. From Proposition 1.2, it follows that

$$\text{haus}_\rho(f, g) \leq \text{haus}_{9\rho}(f, f_\lambda) + \text{haus}_{9\rho}(f_\lambda, g_\lambda) + \text{haus}_{9\rho}(g_\lambda, g).$$

A direct application of the preceding lemmas yields the upper bound. \square

The question of the optimality of the bounds obtained in Theorems 3.4 and 3.7 is important in the derivation of the conditioning number to be associated with a nonlinear optimization problem. This is under investigation by Attouch, Azé, and Peralba [2].

As mentioned in the Introduction, all the results obtained in this section have their counterpart for sets; note however that the constants obtained here may not always be the best possible. Let us consider the case when C, D are nonempty subsets of X . Let $f = \delta_C$ and $g = \delta_D$ be the indicator functions of C and D . Then $\text{haus}_\rho(\delta_C, \delta_D) = \text{haus}_\rho(C, D)$. If $p = 1$, then $(\delta_C)_\lambda = \lambda^{-1} d(\cdot, C)$, and

$$d_{\lambda, \rho}(C, D) = \lambda^{-1} \sup_{\|x\| \leq \rho} |d(x, C) - d(x, D)|.$$

With $\alpha_0 = \alpha_1 = 0$ and

$$\lambda = \left[(18\rho)^{-1} \sup_{\|x\| \leq 9\rho} |d(x, C) - d(x, D)| \right]^{1/2},$$

we obtain the following corollary of Theorem 3.7:

Corollary 3.8. *Suppose C, D are nonempty subsets of a normed linear space X . Then*

$$\text{haus}_\rho(C, D) \leq 2 \left[(18\rho) \sup_{\|x\| \leq 9\rho} |d(x, C) - d(x, D)| \right]^{1/2}.$$

We conclude this section with some remarks concerning the distance

$$d_{\lambda, \rho}^\#(f, g) = \sup_{\|x\| \leq \rho, \|v\|_* \leq \rho} |f_\lambda^\#(x, v) - g_\lambda^\#(x, v)|,$$

where

$$\begin{aligned} f_{\lambda}^{\#}(x, v) &= \inf_{u \in X} \left[f(u) - \langle v, u \rangle + \frac{1}{p\lambda} \|x - u\|^p \right] + \langle v, x \rangle \\ &= \left((f - \langle v, \cdot \rangle + \frac{1}{\lambda p} \|\cdot\|^p)(x) + \langle v, x \rangle \right). \end{aligned}$$

In [7], we introduced these distances to extend some of the results obtained in [6] for $d_{\lambda, \rho}$ when X is a Hilbert space to the situation when X is a reflexive Banach space. To begin with, note that

$$d_{\lambda, \rho}^{\#}(f, g) \geq \sup_{\|x\| \leq \rho} |f_{\lambda}^{\#}(x, 0) - g_{\lambda}^{\#}(x, 0)| = d_{\lambda, \rho}(f, g),$$

and thus from Theorem 3.7 it follows that with the same conditions on f, g, ρ , and λ ,

$$\text{haus}_{\rho}(f, g) \leq d_{\lambda, 9\rho}^{\#}(f, g) + 2(\lambda p)^{1/p} \left[\frac{\alpha'(9\rho)^p + 9\rho + \alpha_1}{1 - p\lambda\alpha'} \right]^{1/p},$$

where $\alpha' = \alpha_0 2^{p-1}$. On the other hand, since $f_{\lambda}^{\#}(x, v) = (f - \langle v, \cdot \rangle)_{\lambda}(x) + \langle v, x \rangle$, it follows that for v fixed the properties of $f_{\lambda}^{\#}(\cdot, v)$ are essentially the same as those of f_{λ} (cf. Lemmas 3.1 and 3.2). Moreover, the same arguments as those in the second proof of Theorem 3.4 (in the Appendix) show that for any $\varepsilon > 0$, for $\|x\| \leq \rho$, and $\|v\|_* \leq \rho$

$$|f_{\lambda}^{\#}(x, v) - g_{\lambda}^{\#}(x, v)| \leq \eta_{\varepsilon} [1 + \rho + \lambda^{-1}(\rho + \gamma + \eta_{\varepsilon})^{p-1}],$$

where η_{ε} and $\gamma = \gamma(\lambda, \rho)$ are the same quantities as those that appear in that proof. Hence,

$$d_{\lambda, \rho}^{\#}(f, g) \leq \beta^{\#}(\lambda, \rho) \text{haus}_{\gamma(\lambda, \rho)}(f, g)$$

with $\beta^{\#}$, the constant β calculated in the proof of Theorem 3.4 plus ρ . We summarize this in the next statement.

Theorem 3.9. *Suppose X is a normed linear space, and f and g are proper, extended real valued functions defined on X such that for given $p \geq 1$, and some $\alpha_0 \geq 0$, $\alpha_1 \in \mathbb{R}$,*

$$f + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0 \quad \text{and} \quad g + \alpha_0 \|\cdot\|^p + \alpha_1 \geq 0.$$

Then for $0 < \lambda < (\alpha_0 p)^{-1} 2^{1-p}$ and $\rho > \max[d(0, \text{epi } f), d(0, \text{epi } g)]$, there exist constants $\beta^{\#}, \gamma, \kappa$, that depend on λ and ρ , such that

$$d_{\lambda, \rho}^{\#}(f, g) \leq \beta^{\#} \text{haus}_{\gamma}(f, g) \leq \beta^{\#} d_{\lambda, 9\rho}^{\#}(f, g) + \kappa,$$

where for fixed ρ, κ can be made arbitrarily small by letting λ go to 0.

4. THE EPI-DISTANCE TOPOLOGY

We limit ourselves to a few basic facts about the topology induced by the pseudo-distances $\{\text{haus}_{\rho}, \rho \geq 0\}$ on the space of extended real functions (for

more about this consult [5]). Our major concern is the relationship between the epi-distance topology and that of epi-convergence. We know that epi-convergence provides the natural conditions, minimal in some sense, under which one can guarantee the convergence of the optimal solutions (see, in particular, [20, §3; 1, §2.2]). Index sets will generally be denoted by N ($N = \mathbb{N}$ in the case of a sequence) with \mathcal{H} a filter a N (\mathcal{H} is the Fréchet filter in the case of a sequence).

Definition 4.1. Let $\overline{\mathbb{R}}^X$ be the space of extended real valued functions defined on the normed linear space X . The topology on $\overline{\mathbb{R}}^X$ generated by the pseudo-distances $\{\text{haus}_\rho, \rho \geq 0\}$ is called the *epi-distance topology*. In other words, for a filtered family $\{f^\nu, \nu \in N\}$ with N an arbitrary (but filtered) index set,

$$f = (\text{epi-dist})\text{-}\lim_{\nu} f^\nu \Leftrightarrow \lim_{\nu} \text{haus}_\rho(f^\nu, f) = 0 \quad \forall \rho \geq 0.$$

Let us begin by observing that the epi-distance topology only depends on the topology of the underlying space X , not on the specific metric that generates this topology. To be convinced of this, it suffices to return to §1, in particular Definition 1.1, and observe that the excess of a set C_ρ on a set D calculated with a specific norm can always be bounded (below or above) by the excess of C_{ρ_1} on D calculated with another equivalent norm for some $\rho_1 > 0$.

We begin by showing that in finite dimension, a collection of functions epi-converges if and only if it converges with respect to the epi-distance. Recall that $\{f^\nu: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}, \nu \in N\}$, a filtered family of functions, is said to *epi-converge* to f if for all $x \in \mathbb{R}^n$:

$$\begin{aligned} &\text{whenever } x^\nu \rightarrow x, \liminf_{\nu} f^\nu(x^\nu) \geq f(x); \text{ and} \\ &\text{there exists } x^\nu \rightarrow x \text{ such that } \limsup_{\nu} f^\nu(x^\nu) \leq f(x). \end{aligned}$$

Theorem 4.2. Suppose $X (= \mathbb{R}^n)$ is finite dimensional. Then the epi-distance topology is the epi-topology, i.e., the topology of epi-convergence.

The proof is a direct consequence of the lemma that follows, the reverse implication is immediate. We extend Theorem 2.2v of Salinetti and Wets [22] to the case of an arbitrary filtered index set. Recall that a filtered family $\{f^\nu, \nu \in N\}$ epi-converges to f if and only if

$$\text{epi } f = \limsup_{\nu} (\text{epi } f^\nu) = \liminf_{\nu} (\text{epi } f^\nu),$$

where, for a family of sets $\{C_\nu \subset X, \nu \in N\}$ with index set N filtered by \mathcal{H} ,

$$\limsup_{\nu} C_\nu = \{x | \forall (Q \in \mathcal{N}(x), H \in \mathcal{H}), \exists \nu \in H \text{ s.t. } C_\nu \cap Q \neq \emptyset\},$$

$$\liminf_{\nu} C_\nu = \{x | \forall Q \in \mathcal{N}(x), \exists H \in \mathcal{H} \text{ s.t. } \forall \nu \in H, C_\nu \cap Q \neq \emptyset\};$$

here $\mathcal{N}(x)$ denotes the neighborhood system of x with respect to the topology on X .

Lemma 4.3. *Let $X = \mathbb{R}^n$, and $\{C_\nu, \nu \in N\}$ be a family of subsets of X filtered by \mathcal{H} . Then, for all $\rho > 0$,*

$$(4.1) \quad \lim_{\nu} e((C_\nu)_\rho, \limsup C_\nu) = 0,$$

$$(4.2) \quad \lim_{\nu} e((\liminf C_\nu)_\rho, C_\nu) = 0.$$

If $C = \liminf_{\nu} C = \limsup_{\nu} C_\nu$, then $\lim_{\nu} \text{haus}_{\rho}(C_\nu, C) = 0$ for $\rho > 0$.

Proof. Let $LS = \limsup_{\nu} C_\nu$, and $LI = \liminf_{\nu} C_\nu$. There is nothing to prove if $LS = \emptyset$, since then, for any $\rho > 0$, there always exists $H \in \mathcal{H}$ such that $(C_\nu)_\rho = \emptyset$ for all $\nu \in H$. Let us thus assume that $LS \neq \emptyset$. If (4.1) does not hold, there exist $\varepsilon > 0$ and $H \in \mathcal{H}^\#$ (the grill of \mathcal{H}) such that for all $\nu \in H$, $e((C_\nu)_\rho, LS) > \varepsilon$, or equivalently for ν in H , there exists $y^\nu \in (C_\nu)_\rho$ such that $d(y^\nu, LS) > \varepsilon$. The collection $\{y^\nu, \nu \in H\} \subset \rho\mathbb{B}$ admits at least one cluster point, say $\bar{y} \in \rho\mathbb{B}$, which also belongs to LS . For this \bar{y} , we have that

$$\lim_{y^\nu \rightarrow \bar{y}} d(y^\nu, LS) = d(\bar{y}, LS) \geq \varepsilon > 0,$$

which of course would contradict the fact that $\bar{y} \in (LS)_\rho$.

Again if $(LI)_\rho = \emptyset$, there is nothing to prove because then $e((LI)_\rho, C_\nu) = 0$ for whatever C_ν . Otherwise, simply observe that $(LI)_\rho \subset LI$, that $e(C_\rho, D) \leq e(C, D)$, and that $\lim_{\nu} e(LI, C_\nu) = 0$ as follows from the definition of the \liminf of a collection of sets. \square

Let us now turn our attention to the infinite-dimensional case, more exactly the case when X is a reflexive Banach space, and epi-limits are defined in terms of Mosco-convergence, i.e., epi-convergence with respect to both the strong and the weak topology on X . Let $\{f^\nu: X \rightarrow \overline{\mathbb{R}}, \nu \in \mathbb{N}\}$ be a sequence of functions. We say that f is the *Mosco-epi-limit* of this sequence, if for all x in X :

whenever $\{x^\nu, \nu \in \mathbb{N}\}$ converges weakly to x , then

$$\liminf_{\nu} f^\nu(x^\nu) \geq f(x);$$

and there exists $\{x^\nu, \nu \in \mathbb{N}\}$ converging strongly to x such that

$$\limsup_{\nu} f^\nu(x^\nu) \leq f(x).$$

Since every Mosco-epi-limit is necessarily weakly lower semicontinuous in infinite dimensions, we are naturally led to focus our attention to the subspace of convex functions. It is then rather easy to see that the convergence of the epi-distances implies Mosco-epi-convergence. Actually, in this setting, the epi-distance topology is strictly finer than the Mosco-epi-topology. We demonstrate all of this in what follows. Also, in the context provided by the important applications of epi-convergence in infinite dimensions, whenever a sequence Mosco-epi-converges to f it also converges with respect to the epi-distance topology.

To begin with let us record an important consequence of Theorems 3.4, 3.7, and 3.9.

Theorem 4.4. *The topology induced on the space of functions $\overline{\mathbb{R}}^X$ defined on the normed linear space X by the pseudo-distance $\{d_{\lambda,\rho}; \lambda > 0, \rho \geq 0\}$ or $\{d_{\lambda,\rho}^\#; \lambda > 0, \rho \geq 0\}$ is the epi-distance topology.*

In the Hilbert case, we know of one more collection of pseudo-distances $\{d_{\lambda,\rho}^J; \lambda > 0, \rho \geq 0\}$ that induces the same topology on the space of proper lower semicontinuous functions on X . This follows from the preceding theorem and [6, Theorem 2.33]. The distance $d_{\lambda,\rho}^J$ is computed as the supremum on ρ -balls of the distance between the resolvents of the Moreau-Yosida approximates of parameter λ . This equivalence is exploited in the proofs of Theorem 5.2 and Proposition 5.4.

In view of this, there appears to be two important topologies that can be defined on the space of proper lower semicontinuous convex functions defined on a reflexive Banach space: the Mosco-epi-topology and the epi-distance topology. The question of knowing when they are equivalent goes begging. One verifies readily that the Mosco-epi-topology is coarser. The example below shows that sometimes it is strictly coarser.

Proposition 4.5. *Suppose X is a reflexive Banach space, and $\{f; f^\nu, \nu \in \mathbb{N}\}$ is a collection of proper, extended real valued, lower semicontinuous, convex functions defined on X . Whenever*

$$\lim_{\nu \rightarrow \infty} \text{haus}_\rho(f, f^\nu) = 0$$

for all ρ sufficiently large, then

$$f = \text{Mosco-epi-lim}_{\nu \rightarrow \infty} f^\nu.$$

Proof. Simply use Theorem 3.4 combined with Theorem 8 of [7]. \square

Example 4.6. Let X be the Hilbert space $L^2(\Omega, \mathbb{R})$,

$$f^\nu(u) = \frac{1}{2} \int_\Omega a_\nu(x) u^2(x) dx, \quad \nu \in \mathbb{N},$$

and

$$f(u) = \frac{1}{2} \int_\Omega a(x) u^2(x) dx.$$

We consider the Moreau-Yosida approximates

$$\begin{aligned} f_\lambda^\nu(u) &= \left(f + \frac{1}{2\lambda} \|\cdot\|^2 \right)(u) \\ &= \frac{1}{2} \int_\Omega \frac{a_\nu}{(1 + \lambda a_\nu)^2} u^2 dx + \frac{1}{2\lambda} \int_\Omega u^2 \left(1 - \frac{1}{1 + \lambda a_\nu} \right)^2 dx \\ &= \frac{1}{2} \int_\Omega \frac{a_\nu}{1 + \lambda a_\nu} u^2 dx. \end{aligned}$$

Hence,

$$\begin{aligned} d_{\lambda,1}(f^\nu, f) &= \sup_{\|u\|_{L^2} \leq 1} |f_\lambda^\nu(u) - f_\lambda(u)| \\ &= \sup_{\|v\|_{L^1} \leq 1} \int_\Omega \left| \frac{a}{1+\lambda a} - \frac{a_\nu}{1+\lambda a_\nu} \right| v \, dx \\ &= \left\| \frac{a}{1+\lambda a} - \frac{a_\nu}{1+\lambda a_\nu} \right\|_{L^\infty}. \end{aligned}$$

Now, take $\Omega = [0, 1]$, $a_\nu(x) = x^{1/\nu}$, and $a(x) = 1$. Then

$$\left\| \frac{1}{1+\lambda} - \frac{1}{x^{-1/\nu} + \lambda} \right\|_\infty = \frac{1}{1+\lambda} = d_{\lambda,1}(f^\nu, f),$$

which does not go to 0. Thus, the f^ν do not converge in the epi-distance topology to f , but they do Mosco-epi-converge. Simply observe that for all $\lambda > 0$, the sequence $\{f_\lambda^\nu, \nu \in \mathbb{N}\}$ is increasing, and pointwise converges to f_λ , which implies Mosco-epi-convergence (see [1, Theorem 3.26]). \square

Usually, however, one is in the situation covered by the next theorem.

Theorem 4.7. *Suppose X and H are two Hilbert space and $X \hookrightarrow H$ is a continuous compact embedding. Then for any collection $\{f; f^\nu, \nu \in \mathbb{N}\}$ of proper, equi-coercive, lower semicontinuous convex functions defined on X , the following statements are equivalent:*

- (i) $f = \text{Mosco-epi-lim}_{\nu \rightarrow \infty} f^\nu$ on H ;
- (ii) for all ρ sufficiently large, $\lim_{\nu \rightarrow \infty} \text{haus}_\rho(f, f^\nu) = 0$;

where the epi-distance is defined in terms of the norm on H . (The collection $\{f^\nu, \nu \in \mathbb{N}\}$ is equi-coercive if there exists a function $\theta: \mathbb{R}_+ \rightarrow [0, \infty)$ with $\lim_{t \rightarrow \infty} \theta(t) = \infty$ such that for all $\nu \in \mathbb{N}$, $f^\nu(x) \geq \theta(\|x\|)$ for all $x \in X$.)

Proof. The assertion follows from [6, Theorem 2.55] and Theorems 3.4 and 3.7. \square

We note that the distance $d_{\lambda,\rho}$ and $d_{\lambda,\rho}^\#$ have been defined here in terms of epigraphical regularizations obtained by taking the epigraphical sum with a polynomial kernel. One can reasonably conjecture that distances defined by epigraphical regularization with respect to a much wider class of kernels are going to be equivalent to the epi-distance. A complete description would be useful; it is an open question.

By relying on the relation between $d_{\lambda,\rho}^\#(f, g)$ and $d_{\lambda,\rho}^\#(f^*, g^*)$ derived in [7, Theorem 5], the next result is obtained as an immediate consequence of Theorem 3.9.

Theorem 4.8. *When X is a reflexive Banach space, the Legendre-Fenchel transform*

$$f \mapsto f^*: \text{plc-fcn}(X) \rightarrow \text{plc-fcn}(X^*)$$

is continuous for the epi-distance topology, where $\text{plc-fcn}(X)$ and $\text{plc-fcn}(X^*)$ are the spaces of proper, extended real valued, lower semicontinuous convex functions defined on X and X^* . In fact, it is uniformly continuous for the uniform structure associated with the pseudo-distance $\{\text{haus}_\rho, \rho \geq 0\}$.

Of course, the epi-distance topology is metrizable. Simply use the pseudo-distances $\{d_{\lambda,\rho}, \lambda > 0, \rho \geq 0\}$ to construct the metric in the standard way. The next result shows that it is also complete, under some restrictions.

Proposition 4.9. *The space of extended real valued functions defined on a normed linear space X equipped with the uniform structure generated by the pseudo-distances $\{\text{haus}_\rho; \rho \geq 0\}$ is complete in the two following situations:*

- (i) X is a finite-dimensional space,
- (ii) X is a reflexive Banach space and the functions are lower semicontinuous, convex and equi-proper, by which we shall mean that all functions are larger than $-\infty$ and there is at least one point at which all are finite valued.

Proof. Let $\{f^\nu; \nu \in \mathbb{N}\}$ be a Cauchy sequence, i.e., for all $\rho > 0$, $\text{haus}_\rho(f^\nu, f^{\nu'}) \rightarrow 0$ as $\nu, \nu' \rightarrow \infty$. From Theorems 3.4 and 3.7, this is equivalent to:

$$\forall \rho > 0, \forall \lambda > 0, d_{\lambda,\rho}(f^\nu, f^{\nu'}) \rightarrow 0 \quad \text{as } \nu, \nu' \rightarrow \infty,$$

where $d_{\lambda,\rho}(f^\nu, f^{\nu'}) = \sup\{|(f^\nu)_\lambda(x) - (f^{\nu'})_\lambda(x)| \mid \|x\| \leq \rho\}$ and $(f^\nu)_\lambda$ is computed for some kernel k of the form $k(\cdot) = \frac{1}{p}\|\cdot\|^p$. We choose to work with $k = \frac{1}{2}\|\cdot\|^2$ to simplify the calculations. Hence, for every $\rho \geq 0$ and $\lambda > 0$, $\{(f^\nu)_\lambda; \nu \in \mathbb{N}\}$ is on $\rho\mathbb{B}$, is a Cauchy sequence for the distance of uniform convergence. Therefore, for every $\lambda > 0$, there exists some function f^λ such that for all $\rho \geq 0$,

$$(f^\nu)_\lambda \rightarrow f^\lambda \quad \text{uniformly on } \rho\mathbb{B}.$$

The crucial step is to show that the family $\{f^\lambda; \lambda > 0\}$ is the epigraphical regularization of a certain function f . If such a function exists it is necessarily given by the formula

$$f := \sup_{\lambda > 0} f^\lambda.$$

Let us compute $(f)_\mu = f +^e (1/2\mu)\|\cdot\|^2$ for $\mu > 0$ and prove that in case (i) or (ii) the equality

$$(f)_\mu = f^\mu$$

holds, which will clearly imply the assertion. We first observe that given any extended real valued function g on X , X being only assumed to be a normed linear space, the epigraphical regularizations of g for various indices are connected by the resolvent equation [6, (2.5)]: $\forall \lambda, \mu > 0, (g_\lambda)_\mu = g_{\lambda+\mu}$. We apply

this with $g = f^\nu$, and pass to the limit as ν goes to ∞ . Noticing that

$$\begin{aligned} ((f^\nu)_\lambda)_\mu(x) &= \inf_{u \in X} [(f^\nu)_\lambda(u) + (1/2\mu)\|x - u\|^2] \\ &= \inf_{u \in \rho_0 \mathbb{B}} [(f^\nu)_\lambda(u) + (1/2\mu)\|x - u\|^2] \end{aligned}$$

for some $\rho_0 \geq 0$ independent of ν , as follows from the uniform convergence of $(f^\nu)_\lambda$ to f^λ on bounded balls in X , we have that $\forall x$, $(f^\lambda)_\mu(x) = \lim_{\nu \rightarrow \infty} ((f^\nu)_\lambda)_\mu(x)$. Since

$$((f^\nu)_\lambda)_\mu = (f^\nu)_{\lambda+\mu} \rightarrow f^{\lambda+\mu} \quad \text{as } \nu \rightarrow \infty,$$

we can conclude that for all $\lambda, \mu > 0$, $(f^\lambda)_\mu = f^{\lambda+\mu}$. Given $\mu > 0$, let us take the supremum with respect to $\lambda > 0$ in this formula. Clearly $\theta \mapsto f^\theta(x)$ is an increasing locally Lipschitz function from \mathbb{R}_+ into \mathbb{R} . Hence, $\sup_{\lambda > 0} (f^\lambda)_\mu = f^\mu$. The only assertion that remains to be established is that $\sup_{\lambda > 0} (f^\lambda)_\mu = f_\mu$. Observing that f^λ increases to f as $\lambda \downarrow 0$, we are in the following situation: given $f^\nu \uparrow f$ does

$$\inf_{u \in X} \left\{ f^\nu(u) + \frac{1}{2\mu} \|x - u\|^2 \right\} \uparrow \inf_{u \in X} \left\{ f(u) + \frac{1}{2\mu} \|x - u\|^2 \right\}?$$

This is clearly satisfied in situations (i) and (ii). In case (ii) just note that the sequence $\{f^\nu; \nu \in \mathbb{N}\}$ Mosco-epi-converges to f and that the set of minimizers of the above expressions is clearly bounded and thus relatively weakly compact. \square

5. MAXIMAL MONOTONE OPERATORS AND SUBGRADIENT MAPPINGS

As for functions, we are interested in a notion of distance between operators that is consistent with a notion of convergence that induces the convergence of the solutions of systems of equations or inclusions defined by these operators. We are referring to *graph-convergence*, i.e., convergence of the graphs (see [15] for recent developments). Here we restrict our attention to maximal monotone operators defined on a Hilbert space X , but the definition that we introduce can be utilized in a more general context, of course. Let

$$\text{gph } A := \{(x, y) \in X \times X | y \in A(x)\}$$

denote the *graph* of the operator $A: X \rightrightarrows X$, possibly multivalued (set-valued) and not necessarily defined everywhere. The operator A is *maximal monotone* if it is monotone, i.e.,

$$\langle y_2 - y_1, x_2 - x_1 \rangle \geq 0 \quad \forall (x_1, y_1) \in \text{gph } A, (x_2, y_2) \in \text{gph } A,$$

and $\text{gph } A$ is a maximal (with respect to inclusion) among all operators that satisfy the preceding condition. From Minty's theorem, we know that this is the case if and only if $\text{rge}(I + A)$, the range of the operator $I + A$, is the whole

space. The *resolvent* of the parameter $\lambda > 0$ is the operator $J_\lambda^A := (I + \lambda A)^{-1}$ defined on the (effective) domain of A , and the *Yosida approximate* of A is the operator $A_\lambda = \lambda^{-1}(I - J_\lambda^A)$.

Let A and B be two maximal monotone operators defined on a Hilbert space X . The ρ -graph distance between A and B , denoted $\text{haus}_\rho(A, B)$, is the ρ -Hausdorff distance between their graphs, i.e.,

$$\text{haus}_\rho(A, B) := \sup[e((\text{gph } A)_\rho, \text{gph } B), e((\text{gph } B)_\rho, \text{gph } A)].$$

We are going to show that these distances are related to the distances $d_{\lambda, \rho}^J$ that measure the distance between operators in terms of their associated resolvents. For all $\lambda > 0$ and $\rho \geq 0$, let

$$d_{\lambda, \rho}^J(A, B) := \sup_{\|x\| \leq \rho} \|J_\lambda^A x - J_\lambda^B x\|$$

denote the “ $d_{\lambda, \rho}^J$ ”-distance between the two operators A and B .

We begin by deriving a bound for the $d_{\lambda, \rho}^J$ in terms of the $\text{haus}_{\rho'}(A, B)$ for some $\rho' > 0$. Let us assume $\text{gph } A \cap \rho_0 \mathbb{B} \neq \emptyset$ and $\text{gph } B \cap \rho_0 \mathbb{B} \neq \emptyset$. Pick $\rho > \rho_0$ and some x in $\rho \mathbb{B}$. From the definition of J_λ^A , it follows that $(J_\lambda^A x, A_\lambda x) \in \text{gph } A$ for $(x_0, y_0) \in \text{gph } A \cap \rho_0 \mathbb{B}$ we have that $x_0 = J_\lambda^A(x_0 + \lambda y_0)$, and because the resolvent is a contraction

$$\|J_\lambda^A x - x_0\| \leq \|x - (x_0 + \lambda y_0)\| \leq \rho + \rho_0(1 + \lambda),$$

which in turn implies that $\|J_\lambda^A x\| \leq \rho + \rho_0(2 + \lambda)$. Thus,

$$\|A_\lambda x\| = \lambda^{-1} \|x - J_\lambda^A x\| \leq \lambda^{-1} [2\rho + \rho_0(2 + \lambda)] \leq 2\lambda^{-1}(\rho + \rho_0) + \rho_0.$$

Let $\rho' = \max[\rho + \rho_0(2 + \lambda), 2\lambda^{-1}(\rho + \rho_0) + \rho_0]$. By definition of the ρ -graph distance there exists $(u, v) \in \text{gph } B$ so that

$$\|u - J_\lambda^A x\| \leq \text{haus}_{\rho'}(A, B), \quad \|v - A_\lambda x\| \leq \text{haus}_{\rho'}(A, B).$$

With $z_\lambda = u + \lambda v$, $u = J_\lambda^B z_\lambda$ and $v = B_\lambda z_\lambda$. Since

$$\begin{aligned} \|z_\lambda - x\| &\leq \|u + \lambda v - x\| = \|u - J_\lambda^A x - \lambda(A_\lambda x - v)\| \\ &\leq \text{haus}_{\rho'}(A, B) + \lambda \text{haus}_{\rho'}(A, B), \end{aligned}$$

and since J_λ^B is a contraction,

$$\|J_\lambda^B z_\lambda - J_\lambda^B x\| \leq \|z_\lambda - x\| \leq (1 + \lambda) \text{haus}_{\rho'}(A, B).$$

Combining this with the fact that u was chosen so that

$$\|u - J_\lambda^A x\| \leq \text{haus}_{\rho'}(A, B),$$

it follows that

$$\|J_\lambda^A x - J_\lambda^B x\| \leq (2 + \lambda) \text{haus}_{\rho'}(A, B).$$

Since x is any point in $\rho\mathbb{B}$, we have $d_{\lambda,\rho}^J(A, B) \leq (2 + \lambda) \text{haus}_{\rho'}(A, B)$.

Next, we are going to show that the ρ -graph distance between two maximal monotone operators A and B is bounded in terms of the distance between their resolvents. We are going to show that with every $(u, v) \in (\text{gph } B)_\rho$ and $\lambda > 0$, we can associate a pair in $\text{gph } A$ such that their distance is bounded by $\lambda' d_{\lambda,(1+\lambda)\rho}^J(A, B)$, where $\lambda' = \max[1, \lambda^{-1}]$. Since this would imply that $e((\text{gph } B)_\rho, \text{gph } A) \leq \lambda' d_{\lambda,(1+\lambda)\rho}^J(A, B)$, a symmetric argument with the roles of A and B interchanged would enable us to assert that

$$\text{haus}_\rho(A, B) \leq \lambda' d_{\lambda,(1+\lambda)\rho}^J(A, B).$$

So, let $(u, v) \in (\text{gph } B)_\rho$ and define $z_\lambda = u + \lambda v$. We again have that $u = J_\lambda^B z_\lambda$ and $v = B_\lambda z_\lambda$. Since $\|z_\lambda\| \leq (1 + \lambda)\rho$, $\|J_\lambda^B z_\lambda - J_\lambda^A z_\lambda\| \leq d_{\lambda,(1+\lambda)\rho}^J(A, B)$. Also, $(J_\lambda^A z_\lambda, AJ_\lambda^A z_\lambda)$ clearly belongs to $\text{gph } A$, and thus to prove the assertion it will suffice to show that $\|AJ_\lambda^A z_\lambda - v\| \leq \lambda^{-1} d_{\lambda,(1+\lambda)\rho}^J(A, B)$. This follows from the simple identity

$$(I + \lambda A)J_\lambda^A z_\lambda = z_\lambda = J_\lambda^B z_\lambda + \lambda v,$$

which tells us that

$$\|AJ_\lambda^A z_\lambda - v\| = \lambda^{-1} \|J_\lambda^B z_\lambda - J_\lambda^A z_\lambda\|,$$

which with the above yields the desired inequality.

The next theorem collects the results of the preceding argument.

Theorem 5.1. *Let X be a Hilbert space and A, B two maximal monotone operators defined on X . Suppose that for some $\rho_0 > 0$, the closed ball $\rho_0\mathbb{B}$ meets both $\text{gph } A$ and $\text{gph } B$. Then for all $\rho \geq \rho_0$ and $\lambda > 0$, we have that*

$$d_{\lambda,\rho}^J(A, B) \leq (2 + \lambda) \text{haus}_{\rho'}(A, B),$$

and

$$\text{haus}_\rho(A, B) \leq \lambda' d_{\lambda,(1+\lambda)\rho}^J(A, B),$$

where $\rho' = \max[\rho + \rho_0(2 + \lambda), 2\lambda^{-1}(\rho + \rho_0) + \rho_0]$ and $\lambda' = \max[1, \lambda^{-1}]$.

Because graph-convergence for a sequence of operators (defined on a reflexive Banach space) is equivalent to the strong convergence of the associated resolvents for all $\lambda > 0$ [1, Proposition 3.60], and convergence with respect to the $d_{\lambda,\rho}^J$ -distances for all ρ larger than some $\rho_0 > 0$ implies the strong convergence of the resolvents, the preceding theorem implies among other things that for maximal monotone operators defined on a Hilbert space, convergence with respect to the ρ -graph distances implies graph-convergence.

Because the subgradient mapping (the subdifferential) of a convex function is a maximal monotone operator, we expect a close relationship between the convergence rates (in terms of the epi-distance) of convex functions and the convergence rates (in terms of the ρ -graph distances) of their subdifferentials.

As a matter of fact, various estimates can already be obtained by relying on the preceding theorem, the results derived in [6, §2], and the basic theorems of §3. This means that we would exploit the results obtained for the $d_{\lambda,\rho}^J$ -distance (operators) and $d_{\lambda,\rho}$ -distance (functions), based on regularizations, to derive the relationship between the ρ -graph distance and the epi-distance. Because of its properties, in particular the characterization provided by the Kenmochi conditions (2.1), the epi-distance is relatively easy to calculate or to estimate in most applications. On the other hand, the distances $d_{\lambda,\rho}$ (or $d_{\lambda,\rho}^\#$) are better suited for theoretical investigations; for example, one can show that the Legendre-Fenchel transform is an isometry for those distances [6]. One of the major consequences of Theorems 3.4, 3.7, and 3.9 is that they give us the flexibility to use either one in our calculation. The proof of our next result illustrates this point. By ∂f we denote the *subgradient mapping*, also called the subdifferential, associated with a convex function f [14]. Let f be a proper convex function with f_λ its Moreau-Yosida approximate of index $\lambda > 0$. The coercivity of $\|\cdot\|^2$ guarantees the existence of a unique point

$$J_\lambda^f(x) := J_\lambda^{\partial f}(x) = (I + \lambda \partial f)^{-1}(x),$$

i.e., the *resolvent* (of parameter λ) of the operator ∂f at x , such that

$$J_\lambda^f(x) = \operatorname{argmin}_u \left[f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right].$$

The function $x \mapsto J_1^f(x)$ is sometimes called the *proximal map*. The pseudo-distance $d_{\lambda,\rho}^J(f, g)$ between two (convex) functions f and g based on the distance between resolvents (introduced in [5]) is the same as the $d_{\lambda,\rho}^J$ -distance between the operators $A = \partial f$ and $B = \partial g$. In the convex context, the two expressions $d_{\lambda,\rho}^J(f, g)$ and $d_{\lambda,\rho}^J(\partial f, \partial g)$ refer to the same quantity.

Theorem 5.2. *Suppose X is a Hilbert space, and f and g are proper, extended real valued, lower semicontinuous convex functions defined on X . To any $\rho > \max[d(0, \operatorname{epi} f), d(0, \operatorname{epi} g)]$ there correspond constants γ and κ (that depend on ρ) such that*

$$\operatorname{haus}_\rho(\partial f, \partial g) \leq \kappa [\operatorname{haus}_\gamma(f, g)]^{1/2}.$$

Proof. As suggested above, we are going to use an intermediate result that is available for the distance $d_{1,\rho}$ computed in terms of the Moreau-Yosida approximates of f and g :

$$f_\lambda = f + {}^e \frac{1}{2} \|\cdot\|^2, \quad g_\lambda = g + {}^e \frac{1}{2} \|\cdot\|^2;$$

the kernel $\frac{1}{2} \|\cdot\|^2$ is particularly well adapted to the Hilbert space setting.

Observe that $(x, y) \in \operatorname{gph} \partial f$ implies that $(x + y) \in (I + \partial f)(x)$, and hence

$$x = (I + \partial f)^{-1}(x + y) = J_1^f(x + y), \quad y = (x + y) - J_1^f(x + y).$$

With $z = x + y$, this yields

$$\text{gph } \partial f = \{(J_1^f(z), z - J_1^f(z)) : z \in H\}.$$

Since J_1^f is a contraction, it follows that $\text{gph } \partial f$ is a Lipschitzian manifold (cf. Brezis [14] and Rockafellar [19]). In particular,

$$(\text{gph } \partial f)_\rho \subset \{(J_1^f(z), z - J_1^f(z)) : \|z\| \leq 2\rho\}.$$

Similarly,

$$\text{gph } \partial g = \{(J_1^g(z), z - J_1^g(z)) : z \in H\}.$$

Thus,

$$e((\text{gph } \partial f)_\rho, \text{gph } \partial g) \leq \sup_{\|z\| \leq 2\rho} \|J_1^f(z) - J_1^g(z)\|.$$

Theorem 2.33 of [6] gives us the inequality

$$d_{1,2\rho}^J(f, g) = \sup_{\|z\| \leq 2\rho} \|J_1^f(z) - J_1^g(z)\| \leq \kappa' [d_{1,\gamma'}(f, g)]^{1/2}$$

with $\kappa' = 2\sqrt{2}$ and $\gamma' = 4\rho + \|J_1^f(0)\| + \|J_1^g(0)\|$. In turn this, with Theorem 3.4, yields,

$$e((\text{gph } \partial f)_\rho, \text{gph } \partial g) \leq \kappa [\text{haus}_\gamma(f, g)]^{1/2},$$

where the constants κ and γ depend on ρ and are the same quantities as those that appear in the calculation in the proof of Theorem 3.4; we note that because f and g are proper convex functions there always exist $\alpha_0 \geq 0$ and $\alpha_1 \in \mathbb{R}$ such that $-\alpha_0 \|\cdot\|^2 - \alpha_1$ minorizes f and g . \square

Note that the inequality $e((\text{gph } \partial f)_\rho, \text{gph } \partial g) \leq \sup_{\|z\| \leq 2\rho} \|J_1^f(z) - J_1^g(z)\|$ used in the proof of the theorem also follows directly from Theorem 5.1; more precisely from $\text{haus}_\rho(A, B) \leq \lambda' d_{\lambda, (1+\lambda)\rho}^J(A, B)$, with $\lambda = 1$, $A = \partial f$, and $B = \partial g$.

Remark 5.3. This last theorem improves a result of Schultz [23] obtained when $X = \mathbb{R}^n$, and f and g are the sum of two continuous convex functions with the same indicator function of a closed convex set. Also, note that the *exponent* $1/2$ is *optimal*. Simply consider $X = \mathbb{R}$, $f(t) = \frac{\alpha}{2}|t|$, and $g(t) = \frac{\alpha}{2}|t - \alpha|$ for some $\alpha > 0$. Then, for $\rho > 0$, $\text{haus}_\rho(f, g) = \frac{1}{2}\alpha^2$ and $\text{haus}_\rho(\partial f, \partial g) = \alpha$.

Proposition 5.4. Suppose X is a Hilbert space, and f and g are proper, extended real valued, lower semicontinuous, convex functions defined on X . Then, for all $\rho > 0$ and $\lambda > 0$,

$$d_{\lambda,\rho}(f, g) \leq (2 + \lambda)\lambda^{-1}\rho \text{haus}_\gamma(\partial f, \partial g) + \alpha_\lambda(f, g),$$

where $\gamma = \gamma(\lambda, \rho) := \sup\{\|J_\lambda^f 0\| + \rho; \lambda^{-1}(2\rho + \|J_\lambda^f 0\|)\}$ and $\alpha_\lambda(f, g) := |f_\lambda(0) - g_\lambda(0)|$.

Proof. With the same notations as in Theorem 5.2, let us start from the inequality [6, Proposition 2.30]

$$d_{\lambda,\rho}(f, g) \leq \lambda^{-1}\rho d_{\lambda,\rho}^J(f, g) + \alpha_\lambda(f, g),$$

with $\alpha_\lambda(f, g)$ as defined above, and

$$d_{\lambda, \rho}^J(f, g) = \sup_{\|x\| \leq \rho} \|J_\lambda^f x - J_\lambda^g x\|,$$

$$J_\lambda^f x = \operatorname{argmin}_u \left[f(u) + \frac{1}{2\lambda} \|x - u\|^2 \right].$$

The remainder of the proof shows that $d_{\lambda, \rho}^J(f, g) \leq (2 + \lambda) \operatorname{haus}_\gamma(\partial f, \partial g)$. (Note that an inequality of the same type can be obtained by applying Theorem 5.1 to the operators ∂f and ∂g .)

From the optimality conditions for $J_\lambda^f x$ in the expression above (if necessary see [14, p. 39] for details), $(J_\lambda^f x, \lambda^{-1}(x - J_\lambda^f x)) \in \operatorname{gph} \partial f$. Moreover, assuming that $\|x\| \leq \rho$, from the contraction properties of J_λ^f it follows that

$$\|J_\lambda^f x\| \leq \|J_\lambda^f 0\| + \rho,$$

$$\|\lambda^{-1}(x - J_\lambda^f x)\| \leq \lambda^{-1}|\rho + \|J_\lambda^f 0\| + \rho|.$$

Hence, $(J_\lambda^f x, \lambda^{-1}(x - J_\lambda^f x)) \in (\operatorname{gph} \partial f)_\gamma$, with γ as defined in the statement of the proposition. If $\operatorname{haus}_\gamma(\partial f, \partial g) < \eta$, then there exists $(y, v) \in \operatorname{gph} \partial g$ such that

$$\|y - J_\lambda^f x\| < \eta, \quad \|v - \lambda^{-1}(x - J_\lambda^f x)\| < \eta.$$

When $u = y + \lambda v$ (cf. the proof of Theorem 5.2), $y = J_\lambda^g u$, $v = \lambda^{-1}(u - J_\lambda^g u)$, we have

$$\|J_\lambda^f x - J_\lambda^g u\| \eta, \quad \lambda^{-1}\|(u - x) - (J_\lambda^g u - J_\lambda^f x)\| < \eta.$$

The last two inequalities imply

$$\|u - x\| < \|J_\lambda^g u - J_\lambda^f x\| + \lambda \eta < (1 + \lambda) \eta.$$

We now use: (i) the triangle inequality,

$$\|J_\lambda^f x - J_\lambda^g x\| \leq \|J_\lambda^f x - J_\lambda^g u\| + \|J_\lambda^g u - J_\lambda^g x\|,$$

(ii) the fact that J_λ^g is a contraction to bound $\|J_\lambda^g u - J_\lambda^g x\|$, and (iii) the bounds on $\|u - x\|$ and $\|J_\lambda^f x - J_\lambda^g u\|$ to conclude

$$\|J_\lambda^f x - J_\lambda^g x\| \leq (2 + \lambda) \eta.$$

Taking the supremum over $\rho \mathbb{B}$ yields $d_{\lambda, \rho}^J(f, g) \leq (2 + \lambda) \eta$. \square

Corollary 5.5. *Under the same assumptions as in Proposition 5.4, and with $\alpha_0 \geq 0$ and $\alpha_1 \in \mathbb{R}$ such that $-\alpha_0 \|\cdot\|^2 - \alpha_1$ minorizes f and g , then for all $0 < \lambda < (4\alpha_0)^{-1}$ and $\rho > \max[d(0, \operatorname{epi} f), d(0, \operatorname{epi} g)]$,*

$$\operatorname{haus}_\rho(f, g) \leq 9(2 + \lambda) \lambda^{-1} \rho \operatorname{haus}_\gamma(\partial f, \partial g) + \kappa,$$

where $\kappa = \alpha_\lambda(f, g) + 4\sqrt{\lambda}(162\alpha_0\rho^2 + 9\rho + \alpha_1)^{1/2}(1 - 4\alpha_0\lambda)^{-1/2}$, and α_λ and $\gamma = \gamma(\lambda, 9\rho)$ are the constants defined in Proposition 5.4.

6. EPI-DISTANCE CALCULUS

In most applications of epigraphical analysis one has to consider extended real-valued functions that have been obtained as the result of various operations involving more elementary functions such as: sums, epi-additions, scalar or epi-multiples, suprema, etc. For example, $f = f_0 + \delta_C$, where δ_C is the indicator function of a set C determined by constraints. Or, for an *averaged problem*, $f = \frac{1}{\nu} *^e (f^1 + {}^e \dots + {}^e f^\nu)$, where each f^i may itself result from the composition of other functions; $*^e$ denotes *epigraphical multiplication*: for $\lambda > 0$, $\lambda *^e f = \lambda f(\lambda^{-1} \cdot)$. Here we shall restrict our attention mostly to sums, epi-additions, and epi-multiplications of functions.

Theorem 6.1. *Let X be a normed linear space, and f^1, g^1, f^2, g^2 extended real valued functions defined on X , bounded below by $\beta \in \mathbb{R}$, and satisfying the following conditions: for every $\rho \geq 0$, and $\theta \geq 0$, there exists a constant $\gamma := \gamma(\theta, \rho) \geq 0$ such that*

$$(6.1) \quad \begin{aligned} (f^1(u) + g^1(v) \leq \theta, \|u + v\| \leq \rho) &\Rightarrow (\|u\| \leq \gamma, \|v\| \leq \gamma), \\ (f^2(u) + g^2(v) \leq \theta, \|u + v\| \leq \rho) &\Rightarrow (\|u\| \leq \gamma, \|v\| \leq \gamma). \end{aligned}$$

Then

$$(6.2) \quad \text{haus}_\rho(f^1 + {}^e g^1, f^2 + {}^e g^2) \leq \text{haus}_{\rho'}(f^1, f^2) + \text{haus}_{\rho'}(g^1, g^2)$$

with $\rho' := \max[\gamma(\rho + 2|\beta| + 1); \gamma(\rho + 2|\beta|); \rho + 3|\beta| + 1]$.

Proof. Pick $x \in (\text{dom}(f^1 + {}^e g^1))_\rho$ such that $|(f^1 + {}^e g^1)(x)| \leq \rho$. By the definition of $f^1 + {}^e g^1$, for every $0 < \varepsilon \leq 1$ there exists $u_\varepsilon \in \text{dom } f^1$ and $v_\varepsilon \in \text{dom } g^1$ such that $u_\varepsilon + v_\varepsilon = x$ and

$$(6.3) \quad f^1(u_\varepsilon) + g^1(v_\varepsilon) - \varepsilon \leq (f^1 + {}^e g^1)(x) \leq \rho.$$

It follows that

$$f^1(u_\varepsilon) + g^1(v_\varepsilon) \leq \rho + 1, \quad \|u_\varepsilon + v_\varepsilon\| \leq \rho.$$

From the assumptions, it then follows that $\|u_\varepsilon\| \leq \gamma(\rho + 1, \rho)$ and $\|v_\varepsilon\| \leq \gamma(\rho + 1, \rho)$; moreover, $\beta \leq f^1(u_\varepsilon) \leq \rho + 1 - \beta$. Set $\rho'' := \max[\gamma(\rho + 1, \rho); \rho + 1 + |\beta|]$; then $\|u_\varepsilon\| \leq \rho''$ and $|f^1(u_\varepsilon)| \leq \rho''$. Kenmochi's conditions (Theorem 2.1(a)) then imply the existence of $u_{\varepsilon,2} \in \text{dom } f^2$ such that

$$\begin{aligned} \|u_\varepsilon - u_{\varepsilon,2}\| &\leq \text{haus}_{\rho''}(f^1, f^2) + \varepsilon, \\ f^2(u_{\varepsilon,2}) &\leq f^1(u_\varepsilon) + \text{haus}_{\rho''}(f^1, f^2) + \varepsilon. \end{aligned}$$

Similarly, there exists $v_{\varepsilon,2} \in \text{dom } g^2$ such that

$$\begin{aligned} \|v_{\varepsilon} - v_{\varepsilon,2}\| &\leq \text{haus}_{\rho''}(g^1, g^2) + \varepsilon, \\ g^2(v_{\varepsilon,2}) &\leq g^1(v_{\varepsilon}) + \text{haus}_{\rho''}(g^1, g^2) + \varepsilon. \end{aligned}$$

From these last pairs of inequalities and (6.3), we have

$$f^2(u_{\varepsilon,2}) + g^2(v_{\varepsilon,2}) \leq (f^1 + {}^c g^1)(x) + 3\varepsilon + \text{haus}_{\rho''}(f^1, f^2) + \text{haus}_{\rho''}(g^1, g^2).$$

With $x_{\varepsilon,2} = u_{\varepsilon,2} + v_{\varepsilon,2}$,

$$\begin{aligned} (f^2 + {}^c g^2)(x_{\varepsilon,2}) &\leq f^2(u_{\varepsilon,2}) + g^2(v_{\varepsilon,2}) \\ &\leq (f^1 + {}^c g^1)(x) + 3\varepsilon + \text{haus}_{\rho''}(f^1, f^2) + \text{haus}_{\rho''}(g^1, g^2), \end{aligned}$$

and

$$\begin{aligned} \|x_{\varepsilon,2} - x\| &= \|(u_{\varepsilon,2} + v_{\varepsilon,2}) - (u_{\varepsilon} + v_{\varepsilon})\| \\ &\leq \|u_{\varepsilon,2} - u_{\varepsilon}\| + \|v_{\varepsilon,2} - v_{\varepsilon}\| \\ &\leq \text{haus}_{\rho''}(f^1, f^2) + \text{haus}_{\rho''}(g^1, g^2) + 2\varepsilon. \end{aligned}$$

Now we use these inequalities, and a symmetric pair obtained when the roles of f^1 and f^2 and g^1 and g^2 are interchanged, in conjunction with the Kenmochi conditions (Theorem 2.1(b)) to obtain

$$(6.2) \quad \text{haus}_{\rho}(f^1 + {}^c g^1, f^2 + {}^c g^2) \leq \text{haus}_{\rho'}(f^1, f^2) + \text{haus}_{\rho'}(g^1, g^2)$$

with $\rho' := \max[\gamma(\rho + 2|\beta| + 1); \gamma(\rho + 2|\beta|); \rho + 3|\beta| + 1]$; note that the functions $f^1 + {}^c g^1$ and $f^2 + {}^c g^2$ are bounded below by 2β . \square

Corollary 6.2. *In the same setting as in Theorem 6.1, conditions (6.1) are satisfied if either f^1 or g^1 is coercive, and at the same time, either f^2 or g^2 is coercive.*

Proof. Indeed if $f(u) + g(v) \leq \theta$ and $\|u + v\| \leq \rho$, then $f(u) \leq \theta - \beta$ and hence $\|u\| \leq \kappa^{-1}(\theta - \beta)$, where κ is the coercivity factor, i.e., $f(u) \geq \kappa(\|u\|)$. Moreover, $\|v\| \leq \|u + v\| + \|u\| \leq \rho + \kappa^{-1}(\theta - \beta)$. Now let $\gamma(\theta, \rho) := \rho + \kappa^{-1}(\theta - \beta)$. \square

Theorem 6.1 and its corollary provide us with the means to estimate the convergence rates of epi-additions; it allows us to write down a quantitative version of the convergence result for epi-additions in [8, Proposition 4.2]. Our next task is to derive a version of Theorem 6.1 for the $d_{\lambda,\rho}$ distances. We begin with a general (and very useful) lemma.

Lemma 6.3. *Suppose f and g are extended real-valued functions defined on a normed linear space X , and, for $\lambda > 0$, f_{λ} and g_{λ} are the Moreau-Yosida approximates of parameter λ . Then*

$$(f + {}^c g)_{\lambda} = f_{\lambda} + {}^c g_{\lambda}.$$

Proof. Observe that $(f + {}^e g)_\lambda = (f + {}^e g) + {}^e (2\lambda)^{-1} \|\cdot\|^2$. Because of the convexity of $(2\lambda)^{-1} \|\cdot\|^2$,

$$(2\lambda)^{-1} \|\cdot\|^2 = \frac{1}{2} * {}^e \left(\frac{1}{2\lambda} \|\cdot\|^2 + {}^e \frac{1}{2\lambda} \|\cdot\|^2 \right).$$

Therefore,

$$(f + {}^e g)_\lambda = (f + {}^e g) + {}^e \left(\frac{1}{2\lambda} \|\cdot\|^2 + {}^e \frac{1}{2\lambda} \|\cdot\|^2 \right).$$

Associativity (of the episum) now yields the desired identity:

$$(f + {}^e g)_\lambda = \left(f + {}^e \frac{1}{2\lambda} \|\cdot\|^2 \right) + {}^e \left(g + {}^e \frac{1}{2\lambda} \|\cdot\|^2 \right) = f_\lambda + {}^e g_\lambda. \quad \square$$

Theorem 6.4. *Let X be a normed linear space, and f^1, g^1, f^2, g^2 extended real valued proper functions defined on X , bounded below by $\beta \in \mathbb{R}$, and satisfying the following condition: for every $\rho \geq 0$ and $\theta \geq 0$, there exists a constant $\gamma := \gamma(\theta, \rho) \geq 0$ such that*

$$(6.1) \quad \begin{aligned} (f^1(u) + g^1(v) \leq \theta, \|u + v\| \leq \rho) &\Rightarrow (\|u\| \leq \gamma, \|v\| \leq \gamma), \\ (f^2(u) + g^2(v) \leq \theta, \|u + v\| \leq \rho) &\Rightarrow (\|u\| \leq \gamma, \|v\| \leq \gamma). \end{aligned}$$

Let ρ_0 be large enough so that the epigraphs of all four functions meet the “ball” $\rho_0 \mathbb{B}_X \times [-\rho_0, \rho_0]$. Then, for any $\lambda > 0$ and $\rho \geq \rho_0$, there exists a constant α (depending on λ and ρ) such that

$$(6.4) \quad d_{\lambda, \rho}(f^1 + {}^e g^1, f^2 + {}^e g^2) \leq d_{\lambda, \alpha}(f^1, f^2) + d_{\lambda, \alpha}(g^1, g^2);$$

the constant α is calculated in the proof of the theorem.

Proof. Suppose $\rho \geq \rho_0$, $\lambda > 0$, and $x \in \rho \mathbb{B}$. Let $u := u_{x, \lambda, \varepsilon}$ and $v := v_{x, \lambda, \varepsilon}$ be such that $u + v = x$ and

$$(6.5) \quad (f^1 + {}^e g^1)_\lambda(x) = (f^1_\lambda + {}^e g^1_\lambda)(x) \geq f^1_\lambda(u) + g^1_\lambda(v) - \varepsilon.$$

If we can show that there exists $\alpha = \alpha(\lambda, \rho)$ such that both u and v belong to $\alpha \mathbb{B}$, then from the definition of the pseudo-distances $d_{\lambda, \rho}$, it would follow that

$$f^1_\lambda(u) \geq f^2_\lambda(u) - d_{\lambda, \alpha}(f^1, f^2), \quad g^1_\lambda(v) \geq g^2_\lambda(v) - d_{\lambda, \alpha}(g^1, g^2),$$

which combined with the above would yield

$$\begin{aligned} (f^1 + {}^e g^1)_\lambda(x) &\geq f^2_\lambda(u) + g^2_\lambda(v) - d_{\lambda, \alpha}(f^1, f^2) - d_{\lambda, \alpha}(g^1, g^2) - \varepsilon \\ &\geq (f^2 + {}^e g^2)_\lambda(x) - d_{\lambda, \alpha}(f^1, f^2) - d_{\lambda, \alpha}(g^1, g^2) - \varepsilon \\ &\geq (f^2 + {}^e g^2)_\lambda(x) - d_{\lambda, \alpha}(f^1, f^2) - d_{\lambda, \alpha}(g^1, g^2) - \varepsilon. \end{aligned}$$

Interchanging the roles of f^1 and f^2 and g^1 and g^2 , letting ε go to 0, and taking suprema with respect to $x \in \rho \mathbb{B}$ yields (6.4).

So we only need to prove that there exists an α such that $\|u\| \leq \alpha$ and $\|v\| \leq \alpha$. On one hand, we have that

$$\begin{aligned} (f^1 + {}^e g^1)_\lambda(x) &\leq (f^1 + {}^e g^1)(s) + (2\lambda)^{-1} \|x - s\|^2 \quad \forall s \in X, \\ &\leq f^1(r) + g^1(s) + (2\lambda)^{-1} \|x - (r + s)\|^2 \quad \forall s, r \in X, \end{aligned}$$

and hence

$$(6.6) \quad (f^1 + {}^e g^1)_\lambda(x) \leq 2\rho_0 + (2\lambda)^{-1}(\rho + 2\rho_0)^2.$$

On the other hand, for every $\varepsilon > 0$, there exists $\hat{u} := \hat{u}_{\lambda, \varepsilon}$ and $\hat{v} := \hat{v}_{\lambda, \varepsilon}$ such that

$$\begin{aligned} f_\lambda^1(u) &\geq f^1(\hat{u}) + (2\lambda)^{-1} \|u - \hat{u}\|^2 - \varepsilon, \\ g_\lambda^1(v) &\geq g^1(\hat{v}) + (2\lambda)^{-1} \|v - \hat{v}\|^2 - \varepsilon, \end{aligned}$$

and thus

$$f_\lambda^1(u) + g_\lambda^1(v) \geq f^1(\hat{u}) + g^1(\hat{v}) + (2\lambda)^{-1} \|u - \hat{u}\|^2 + (2\lambda)^{-1} \|v - \hat{v}\|^2 - 2\varepsilon.$$

This with (6.5) and (6.6) yields the inequality

$$2\rho_0 + (2\lambda)^{-1}(\rho + 2\rho_0)^2 \geq f^1(\hat{u}) + g^1(\hat{v}) + 2(\lambda)^{-1} \|u - \hat{u}\|^2 + (2\lambda)^{-1} \|v - \hat{v}\|^2 - 3\varepsilon,$$

which implies that

$$\begin{aligned} \mu &:= 2\rho_0 + (2\lambda)^{-1}(\rho + 2\rho_0)^2 + 3 \geq f^1(\hat{u}) + g^1(\hat{v}), \\ 2\rho_0 + (2\lambda)^{-1}(\rho + 2\rho_0)^2 + 3 - 2\beta &\geq (2\lambda)^{-1}(\|u - \hat{u}\|^2 + \|v - \hat{v}\|^2). \end{aligned}$$

Since

$$\|\hat{u} + \hat{v}\| \leq \|\hat{u} + \hat{v} - (u + v)\| + \|u + v\| \leq \|\hat{u} - u\| + \|\hat{v} - v\| + \rho,$$

with the above this yields

$$\|\hat{u} + \hat{v}\| \leq \rho^\dagger := [4\lambda(2\rho_0 + (2\lambda)^{-1}(\rho + 2\rho_0)^2 + 3 - 2\beta)]^{1/2}.$$

The assumptions (6.1) with what precedes yields the following estimates:

$$\begin{aligned} \|\hat{u}\| &\leq \gamma(\mu, \rho^\dagger), & \|\hat{v}\| &\leq \gamma(\mu, \rho^\dagger), \\ \|u - \hat{u}\| &\leq [2\lambda(\mu - 2\beta)]^{1/2}, & \|v - \hat{v}\| &\leq [2\lambda(\mu - 2\beta)]^{1/2}, \end{aligned}$$

and hence

$$(6.7) \quad \|u\| \leq \gamma(\mu, \rho^\dagger) + [2\lambda(\mu - 2\beta)]^{1/2} := \alpha(\lambda, \rho),$$

and similarly for $\|v\|$. \square

Note that the “constant” α calculated in the proof of the preceding theorem depends continuously on λ , ρ , ρ_0 , and μ ; it tends to ∞ when either $\lambda \downarrow 0$ or $\rho \uparrow \infty$.

Our next result will be a “dual” version of Theorem 6.4 involving the sum of (convex) functions rather than their epi-sum. In what follows X is a Hilbert

space, and $\text{plc-fcn}(X)$ will denote the space of proper, lower semicontinuous convex functions defined on X with values in $\mathbb{R} \cup \{\infty\}$. For $f \in \text{plc-fcn}(X)$, its *conjugate* is the function:

$$f^*(v) = \sup_{x \in X} \{\langle v, x \rangle - f(x)\}.$$

The mapping that associates its conjugate to each member of $\text{plc-fcn}(X)$ also in $\text{plc-fcn}(X)$, is called the *Legendre-Fenchel transform*. In [5, §2], we showed that the Legendre-Fenchel transform is an *isometry* for the $d_{1,\rho}$ pseudo-distances. More generally, for all $f, g \in \text{plc-fcn}(X)$ and all $\lambda > 0$, $\rho \geq 0$,

$$d_{\lambda,\rho}(f, g) = d_{\lambda^{-1}, \rho\lambda^{-1}}(f^*, g^*).$$

Lemma 6.5. *Suppose $f, g \in \text{plc-fcn}(X)$ satisfy the following condition:*

there exists some ρ_0 such that $\rho_0\mathbb{B} \subset \text{dom } f - \text{dom } g$ and for every $x \in \rho_0\mathbb{B}$ there exists $u \in \text{dom } f$ and $v \in \text{dom } g$ such that for some $\alpha := \alpha(\rho_0)$:

$$z = u - v \quad \text{and} \quad \max[\|u\|, \|v\|, f(u), g(v)] \leq \alpha.$$

Then, f^ and g^* satisfy the condition (of Theorem 6.1): for every $\rho \geq 0$ and $\mu \geq 0$, there exists a “constant” $\gamma \geq 0$ such that*

$$(6.8) \quad (f^*(u) + g^*(v) \leq \mu, \|u + v\| \leq \rho) \Rightarrow (\|u\| \leq \gamma, \|v\| \leq \gamma),$$

with the “constant” γ depending continuously on ρ , μ , and α .

Proof. Let us assume that f and g satisfy the assumptions. Pick $\rho \geq 0$, $\mu \geq 0$, and $\bar{u}, \bar{v} \in X$ such that

$$f^*(\bar{u}) + g^*(\bar{v}) \leq \mu, \quad \|\bar{u} + \bar{v}\| \leq \rho.$$

By assumption, for all $z \in \rho_0\mathbb{B}$ there exists $u \in \text{dom } f$ and $v \in \text{dom } g$ such that $z = u - v$. We then observe that

$$\begin{aligned} \langle \bar{u}, z \rangle &= \langle \bar{u}, u - v \rangle = \langle \bar{u}, u \rangle + \langle \bar{v}, v \rangle - \langle \bar{u} + \bar{v}, v \rangle \\ &\leq \langle \bar{u}, v \rangle + \langle \bar{v}, v \rangle + \|\bar{u} + \bar{v}\| \cdot \|v\| \\ &\leq f(u) + f^*(\bar{u}) + g(v) + g^*(\bar{v}) + \|\bar{u} + \bar{v}\| \cdot \|v\| \\ &\leq \mu + f(u) + g(v) + \rho\|v\|. \end{aligned}$$

The condition satisfied by f and g now implies that

$$\|\bar{u}\| \leq \frac{1}{\rho_0}[\mu + (2 + \rho)\alpha] := \gamma(\mu, \rho).$$

A symmetric argument with \bar{v} instead of \bar{u} completes the proof. \square

Theorem 6.6. *Let X be a Hilbert space and $f^1, g^1, f^2, g^2 \in \text{plc-fcn}(X)$. Suppose that both pairs f^1 and g^1 , and f^2 and g^2 satisfy the condition of Lemma 6.5, that*

$$0 \in \text{dom } f^1 \cap \text{dom } g^1, \quad 0 \in \text{dom } f^2 \cap \text{dom } g^2,$$

and that there exists $\rho_0^\# \geq 0$ and x_0 such that $f^1 \geq \langle x_0, \cdot \rangle - \rho_0^\#$, and similarly for g^1 , f^2 and g^2 . Then for any $\lambda > 0$ and $\rho \geq 0$, we have

$$d_{\lambda, \rho}(f^1 + g^1, f^2 + g^2) \leq d_{\lambda, \beta}(f^1, f^2) + d_{\lambda, \beta}(g^1, g^2),$$

with $\beta = \beta(\lambda, \rho)$ calculated in the proof.

Proof. For $i = 1, 2$, set $h^i = (f^i)^*$, $l^i = (g^i)^*$. Since $\rho_0 \mathbb{B} \subset \text{dom } f^1 - \text{dom } f^2$, $(f^1 + g^1)^* = h^1 + {}^e l^1$ (cf. [11]). Moreover, $-f^1(0) = \inf h^1 > -\infty$, and similarly for f^2 , g^1 , and g^2 . From the preceding lemma it follows that h^1 , l^1 , as well as h^2 , l^2 , satisfy (6.8) with $\gamma(\mu, \rho) = \rho_0^{-1}[\mu + (2 + \rho)\alpha(\rho_0)]$. Moreover, $f^1 \geq \langle x_0, \cdot \rangle - \rho_0^\#$ implies that $l^1(x_0) \leq \rho_0^\#$ with $\|x_0\| \leq \rho_0^\#$. Thus, we can apply Theorem 6.4 to obtain

$$d_{\lambda^{-1}, \lambda^{-1}\rho}(h^1 + {}^e l^1, h^2 + {}^e l^2) \leq d_{\lambda^{-1}, \alpha}(h^1, h^2) + d_{\lambda^{-1}, \alpha}(l^1, l^2),$$

where $\alpha := \alpha(\lambda^{-1}, \rho\lambda^{-1})$ is the quantity defined by relation (6.7) (when the first argument is λ^{-1} and the second one $\rho\lambda^{-1}$). From the isometry of the Legendre-Fenchel transform for these pseudo-distances [6], it follows that

$$d_{\lambda, \rho}(f^1 + {}^e g^1, f^2 + {}^e g^2) \leq d_{\lambda, \lambda\alpha}(f^1, f^2) d_{\lambda, \lambda\alpha}(g^1, g^2),$$

and that is the assertion of the theorem, with $\beta = \lambda\alpha(\lambda^{-1}, \rho\lambda^{-1})$; more precisely,

$$\begin{aligned} \mu_0 &= \max[0, f^1(0), f^2(0), g^1(0), g^2(0)], \\ \rho^\# &= \sqrt{4\lambda^{-1}(2\rho_0^\# + \tfrac{1}{2}\lambda(\rho\lambda^{-1} + 2\rho_0^\#)^2 + 3 + 2\mu_0)}, \\ \mu &= 2\rho_0^\# + \tfrac{1}{2}\lambda(\rho\lambda^{-1} + 2\rho_0^\#)^2 + 3, \\ \beta(\lambda, \rho) &= \lambda[\mu + (2 + \rho^\#)\alpha(\rho_0)]/\rho_0^\# + [2\lambda(\mu + 2\mu_0)]^{1/2}, \end{aligned}$$

using here the constants calculated in Theorem 6.4. \square

Remark 6.7. One could relax the assumptions in Lemma 6.5 and Theorem 6.6 by letting the “constant” α also depend on the point z . The conclusion would still hold thanks to the Banach-Steinhaus theorem.

The preceding theorem can be compared to a result of D. Azé [11, Chapitre 6, Proposition 2.3]. A similar result expressed in terms of the epi-distances has recently been obtained by Azé and Penot. We record it here for the convenience of the reader.

Proposition 6.8 [13]. *Let X be a Banach space, and consider functions $(f_i, i = 1, \dots, n)$ and $(g_i, i = 1, \dots, n)$ in $\text{plc-fcn}(X)$. Assume that these functions are minorized by $-\alpha(\|\cdot\| + 1)$ for some $\alpha \geq 0$ and that for some $\gamma \geq 0$ and $\sigma > 0$,*

$$(\sigma\mathbb{B})^n \subset \text{diag } X^n \cap (\gamma\mathbb{B})^n - \prod_{i=1}^n (\text{lev}_\gamma f_i)_\gamma,$$

where “ $-$ ” is the Minkowski operation, $\text{diag } X^n := \{(x, \dots, x) | x \in X\}$, and $\text{lev}_\gamma f_i := \{x | f_i(x) \leq \gamma\}$. Then, whenever $\sum_{i=1}^n \text{haus}_{\rho_1}(f_i, g_i) < \sigma$,

$$\text{haus}_\rho \left(\sum_{i=1}^n f_i, \sum_{i=1}^n g_i \right) \leq \sigma^{-1} (n\gamma + \sigma + \rho) \sum_{i=1}^n \text{haus}_{\rho_1}(f_i, g_i) \quad \forall \rho \geq n\gamma + \sigma,$$

where $\rho_1 = \rho + (n+1)[\alpha(\rho + \sigma + 1) + \sigma]$.

We conclude with a result about epi-multiplication. Recall that if $f: X \rightarrow \overline{\mathbb{R}}$ with X a normed linear space, and $\lambda > 0$, then

$$\lambda \text{ epi } f = \text{epi}(\lambda *^e f), \quad \text{where } (\lambda *^e f)(x) = \lambda f(\lambda^{-1}x).$$

Moreover, $(\lambda *^e f)^* = \lambda f^*$, i.e., $*^e$ is the conjugate operation of the standard scalar multiplication (by positive scalars).

Proposition 6.9. Suppose $f, g: X \rightarrow \mathbb{R} \cup \{\infty\}$ are both minorized by $-\alpha_0 \|\cdot\|^p - \alpha_1$ for some $\alpha_0 > 0$, $\alpha_1 \in \mathbb{R}$, and $1 \leq p < \infty$. Then, for $\lambda > 0$, $\mu > 0$, and every $\rho \geq 0$,

$$\begin{aligned} & \text{haus}_\rho(\lambda *^e f, \mu *^e g) \\ & \leq \max[\lambda', \mu'] \cdot |\lambda - \mu| \rho + \sup\{\mu \text{haus}_{\rho'\lambda'}(f, g); \lambda \text{haus}_{\rho'\mu'}(f, g)\}, \end{aligned}$$

where $\rho' := \rho + \alpha_0 \rho^p + \alpha_1$ and $\lambda' = \lambda^{-1}$, $\mu' = \mu^{-1}$. In particular, if $\mu = \lambda$,

$$\text{haus}_\rho(\lambda *^e f, \lambda *^e g) \leq \lambda \text{haus}_{\rho'\lambda'}(f, g).$$

Proof. Pick $x \in \rho\mathbb{B}$ with $|(\lambda *^e f)(x)| \leq \rho$. Then $|f(\lambda^{-1}x)| \leq \rho\lambda^{-1}$ and $\|\lambda^{-1}x\| \leq \rho\lambda^{-1}$. It follows from Kenmochi's conditions (Theorem 2.1(a)) that for every $\varepsilon > 0$ there exists some $y_\varepsilon \in \text{dom } g$ that satisfies

$$\|y_\varepsilon - \lambda^{-1}x\| \leq \text{haus}_{\rho\lambda^{-1}}(f, g) + \varepsilon, \quad g(y_\varepsilon) \leq f(\lambda^{-1}x) + \text{haus}_{\rho\lambda^{-1}}(f, g) + \varepsilon.$$

From this it follows that

$$\begin{aligned} \mu g \left(\frac{\mu y_\varepsilon}{\mu} \right) & \leq \mu f(\lambda^{-1}x) + \mu \text{haus}_{\rho\lambda^{-1}}(f, g) + \mu\varepsilon \\ & \leq \lambda f(\lambda^{-1}x) + (\mu - \lambda)f(\lambda^{-1}x) + \mu \text{haus}_{\rho\lambda^{-1}}(f, g) + \mu\varepsilon. \end{aligned}$$

Similarly, one obtains

$$\left\| \mu y_\varepsilon - \frac{\mu x}{\lambda} \right\| \leq \mu \text{haus}_{\rho\lambda^{-1}}(f, g) + \mu\varepsilon.$$

Hence,

$$\begin{aligned} (\mu *^e g)(\mu y_\varepsilon) & \leq (\lambda *^e f)(x) + |\mu - \lambda| \rho \lambda^{-1} + \mu \text{haus}_{\rho\lambda^{-1}}(f, g) + \mu\varepsilon, \\ \|\mu y_\varepsilon - x\| & \leq |1 - \mu\lambda^{-1}| \rho + \text{haus}_{\rho\lambda^{-1}}(f, g) + \mu\varepsilon. \end{aligned}$$

The assertion now follows from the Kenmochi's conditions (Theorem 2.1).

□

APPENDIX

Second proof of Theorem 3.4. As in the proof of Theorem 3.4, we only need to consider the case when f and g are proper. Pick $x \in X$, $0 < \lambda < (\alpha_0 p)^{-1} 2^{-1}$, and let us calculate an upper bound for $f_\lambda(x) - g_\lambda(x)$. For $0 < \varepsilon \leq 1$, let u_ε be such that

$$g(u_\varepsilon) + (\lambda p)^{-1} \|x - u_\varepsilon\|^p \leq g_\lambda(x) + \varepsilon,$$

i.e., u_ε attains, up to ε , the infimum in the definition of g_λ . Then

$$(A.1) \quad f_\lambda(x) - g_\lambda(x) \leq \inf_u \{f(u) - g(u_\varepsilon) + (\lambda p)^{-1} (\|x - u\|^p - \|x - u_\varepsilon\|^p)\} + \varepsilon.$$

Let us begin by deriving an estimate for $\|u_\varepsilon\|$. From the minorization of g , it follows that (with $\alpha := \max[\alpha_0, \alpha_1]$).

$$-\alpha(\|u_\varepsilon\|^p + 1) + \frac{1}{\lambda p} \|x - u_\varepsilon\|^p \leq g_\lambda(x) + \varepsilon \leq g(u_0) + (\lambda p)^{-1} \|x - u_0\|^p + \varepsilon,$$

where u_0 is some arbitrary point in $\text{dom } g$. Hence,

$$(\lambda p)^{-1} \|x - u_\varepsilon\|^p \leq \alpha(1 + 2^{p-1} \|u_\varepsilon - x\|^p + 2^{p-1} \|x\|^p) + g(u_0) + (\lambda p)^{-1} \|x - u_0\|^p + \varepsilon,$$

$$\|x - u_\varepsilon\|^p \leq ((\lambda p)^{-1} - \alpha 2^{p-1})^{-1} [\alpha(1 + 2^{p-1} \|x\|^p) + g(u_0) + (\lambda p)^{-1} \|x - u_0\|^p + \varepsilon],$$

and since $\|u_\varepsilon\|^p \leq 2^{p-1} (\|u_\varepsilon - x\|^p + \|x\|^p)$, when $\|x\| \leq \rho$:

$$\begin{aligned} \|u_\varepsilon\|^p &\leq 2^{p-1} [\rho^p + ((\lambda p)^{-1} - \alpha 2^{p-1})^{-1} [\alpha(1 + 2^{p-1} \rho^p) \\ &\quad + g(u_0) + (\lambda p)^{-1} (\|u_0\| + \rho)^p + \varepsilon]]. \end{aligned}$$

This means that $\|u_\varepsilon\|$ is bounded above by a constant that depends on ρ , $\|u_0\|$, $g(u_0)$, α , and λ . We are interested in the dependence on λ and ρ , and express this by writing $\|u_\varepsilon\| \leq \gamma_1(\lambda, \rho)$. Next, we calculate an estimate for $g(u_\varepsilon)$. We have

$$g(u_\varepsilon) \geq -\alpha(1 + \|u_\varepsilon\|^p) \geq -\alpha(1 + \gamma_1(\lambda, \rho)^p).$$

Also

$$g(u_\varepsilon) \leq g_\lambda(x) + \varepsilon \leq g(u_0) + (\lambda p)^{-1} \|x - u_0\|^p + \varepsilon.$$

Hence,

$$|g(u_\varepsilon)| \leq \max[\alpha(1 + \gamma_1(\lambda, \rho)^p); g(u_0) + (\lambda p)^{-1} \|x - u_0\|^p + 1] =: \gamma_2(\lambda, \rho).$$

Thus, with $\gamma(\lambda, \rho) := \max[\gamma_1(\lambda, \rho); \gamma_2(\lambda, \rho)]$, we have that $(u_\varepsilon, g(u_\varepsilon)) \in \mathbb{B}_{X \times \mathbb{R}}(0, \gamma(\lambda, \rho))$. By Theorem 2.1, more precisely by the Kenmochi conditions (2.1), we know that there exists v_ε such that

$$\begin{aligned} \|v_\varepsilon - u_\varepsilon\| &\leq \text{haus}_{\gamma(\lambda, \rho)}(f, g) + \varepsilon =: \eta_\varepsilon \\ f(v_\varepsilon) &\leq g(u_\varepsilon) + \text{haus}_{\gamma(\lambda, \rho)}(f, g) + \varepsilon = g(u_\varepsilon) + \eta_\varepsilon. \end{aligned}$$

From (A.1), it follows that

$$f_\lambda(x) - g_\lambda(x) \leq f(v_\varepsilon) + (\lambda p)^{-1} \|x - v_\varepsilon\|^p - g(u_\varepsilon) - (\lambda p)^{-1} \|x - u_\varepsilon\|^p + \varepsilon,$$

which combined with the preceding inequalities yields

$$\begin{aligned} f_\lambda(x) - g_\lambda(x) &\leq \eta_\varepsilon + \lambda^{-1} \|x - v_\varepsilon\|^{p-1} \|u_\varepsilon - v_\varepsilon\| + \varepsilon, \\ &\leq \eta_\varepsilon [1 + \lambda^{-1} (\rho + \gamma(\lambda, \rho) + \eta_\varepsilon)^{p-1}] + \varepsilon. \end{aligned}$$

Because this holds for all $\varepsilon > 0$, and because a similar inequality holds when the role of f and g are interchanged, we have

$$\begin{aligned} |f_\lambda(x) - g_\lambda(x)| &\leq \text{haus}_{\gamma(\lambda, \rho)}(f, g) [1 + \lambda^{-1} (\rho + \gamma(\lambda, \rho) + \text{haus}_{\gamma(\lambda, \rho)}(f, g)^{p-1})] \\ &\leq \text{haus}_{\gamma(\lambda, \rho)}(f, g) \beta(\lambda, \rho). \end{aligned}$$

This completes the proof, since x is an arbitrary point in $\rho\mathbb{B}$. \square

REFERENCES

1. H. Attouch, *Variational convergence for functions and operators*, Appl. Math. Ser., Pitman, London, 1984.
2. H. Attouch, D. Azé, and J.-C. Peralba, *Comparaison de différentes métriques liées à l'épi-convergence* (in preparation).
3. H. Attouch, D. Azé and R. J.-B. Wets, *On the continuity properties of the partial Legendre-Fenchel transform: convergence of sequences of augmented lagrangian functions*, Moreau-Yosida approximates and subdifferential operators, Optimization Days 1985, Proc. Journées Fermat (J.-B. Hiriart-Urruty, ed.), North-Holland, Amsterdam, 1986.
4. H. Attouch and A. Damlamian, *Problèmes d'évolution dans les Hilbert et applications*, J. Math. Pures Appl. **54** (1975), 53–74.
5. H. Attouch, R. Lucchetti, and R. J.-B. Wets, *The topology of the p -Hausdorff distance*, Ann. Mat. Pura Appl. Ann. Mat. Pura Appl. (to appear).
6. H. Attouch and R. J.-B. Wets, *Isometries for the Legendre-Fenchel transform*, Trans. Amer. Math. Soc. **296** (1986), 33–60.
7. —, *Another isometry for the Legendre-Fenchel transform*, J. Math. Anal. Appl. **131** (1988), 404–411.
8. —, *Epigraphical analysis*, Analyse Non Linéaire (H. Attouch, J.-P. Aubin, F. Clarke, and I. Ekeland, eds.), Gauthier-Villars, Paris, 1989, pp. 73–100.
9. —, *Quantitative stability of variational systems: II. A framework for nonlinear conditioning*, IIASA Working Paper 88-9, Laxenburg, Austria, February 1988.
10. —, *Lipschitzian stability of the ε -approximate solutions in convex optimization*, IIASA Working Paper WP-87-25, Laxenburg, Austria, March 1987.
11. J.-P. Aubin, *Comportement lipschitzien des solutions de problèmes de minimisation convexes*, C. R. Acad. Sci. Paris **295** (1982), 235–238.
12. D. Azé, *Convergence variationnelle et dualité. Applications en calcul des variations et en programmation mathématique*, Thèse de Doctorat, AVAMAC-Perpignan, 1985.
13. D. Azé and J.-P. Penot, *Operations on convergent families of sets and functions*, Technical Report AVAMAC, Perpignan, 1987.
14. H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland, Amsterdam, 1973.
15. V. Chiado Piat, G. Dal Maso, and A. Defranceschi, *G-convergence of monotone operators*, SISSA Ref. 84M, Trieste, July 1988.
16. N. Kenmochi, *The semi-discretization method and time dependent parabolic variational inequalities*, Proc. Japan Acad. Ser. A Math. Sci. **50** (1974), 714–717.
17. J.-J. Moreau, *Intersection of moving convex sets in a normed space*, Math. Scand. **36** (1975), 159–173.

18. U. Mosco, *Convergence of convex sets and of solutions of variational inequalities*, Adv. in Math. **3** (1969).
19. R. T. Rockafellar, *Maximal monotone relations and the second derivative of nonsmooth functions*, Ann. Inst. H. Poincaré Anal. Non Linéaire **2** (1985), 510–585.
20. R. T. Rockafellar and R. J.-B. Wets, *Variational systems, an introduction*, Multifunctions and Integrands: Stochastic Analysis, Approximation and Optimization (G. Salinetti, ed.) Lecture Notes in Math., vol. 1091, Springer-Verlag, Berlin, 1984, pp. 1–54.
21. G. Salinetti and R. J.-B. Wets, *On the convergence of sequence of convex sets in finite dimensions*, SIAM Rev. **21** (1979), 16–33.
22. ———, *On the convergence of closed-valued measurable multifunctions*, Trans. Amer. Math. Soc. **266** (1981), 275–289.
23. R. Schultz, *Estimates for Kuhn-Tucker points of perturbed convex programs*, Technical Report, Humboldt Univ., Berlin, 1986.
24. D. Walkup and R. J.-B. Wets, *Continuity of some convex cone-valued mappings*, Proc. Amer. Math. Soc. **18** (1967), 229–235.
25. R. J.-B. Wets, *Convergence of convex functions, variational inequalities and convex optimization problems*, Variational Inequalities and Complementarity Problems (R. Cottle, F. Giannessi, and J. L. Lions, eds.), Wiley, Chichester, 1980, pp. 405–419.

DÉPARTEMENT DE MATHÉMATIQUES, UNIVERSITÉ DU LANGUEDOC, 34060 MONTPELLIER, FRANCE

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA–DAVIS, DAVIS, CALIFORNIA 95616